HortScience 56(4):481–490. 2021. https://doi.org/10.21273/HORTSCI15607-20

Identification of Phenotypic Characteristics in Three Chemotype Categories in the Genus *Cannabis*

Dan Jin
Department of Biomedical Engineering, University of Alberta, Edmonton, Alberta, Canada; and PBG BioPharma, Leduc, Alberta, Canada

Philippe Henry
Egret Bioscience Ltd., West Kelowna, British Columbia, Canada; and Lighthouse Genomics Inc., Salt Spring Island, British Columbia, Canada

Jacqueline Shan
PBG BioPharma, Leduc, Alberta, Canada

Jie Chen
Department of Biomedical Engineering, University of Alberta, Edmonton, Alberta, Canada; and Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta, Canada

Additional index words. Cannabis, THC, CBD, phenotypic markers, chemotypes

Abstract. **Modern *Cannabis* cultivars are morphologically distinguished by their leaflet shapes (wide for "Indica" and narrow for "Sativa") by users and breeders. However, there are no scientific bases or references for determining the shape of these leaflets. In addition, these two categories contained mostly THC dominant (high THC) cultivars while excluded CBD dominant (high CBD) and intermediate (intermediate level of both THC and CBD) cultivars. This study investigated the phenotypic variation in 21 *Cannabis* cultivars covering three chemical phenotypes, referred to as chemotypes, grown in a commercial greenhouse. Thirty morphological traits were measured in the vegetative, flowering, and harvest stages on live plants and harvested inflorescences. The collected data were subjected to correlation analysis, hierarchical clustering, principal component analysis, and canonical correlation analysis with preassigned chemotypes. Canonical correlation analysis assigned individual plants to their chemotypes with 92.9% accuracy. Significant morphological differences were identified. Traits usable as phenotype markers for CBD dominant cultivars included light-green and narrow leaflets, a greater number of primary and secondary serrations, loose inflorescences, dense and resinous trichomes, and *Botrytis cinerea* resistance. Traits for intermediate cultivars included deep-green and medium-wide leaflets, more primary and secondary serrations, medium compact inflorescences, trichomes that are less dense and less resinous, and *Botrytis cinerea* resistance. Traits for THC dominant cultivars included deep-green and wide leaflets, large and compact inflorescences, dense and resinous trichomes, and *Botrytis cinerea* susceptibility. The results of this study provide a comprehensive profile of morphological traits of modern *Cannabis* cultivars and provides the first such profile for CBD dominant and intermediate cultivars. Additionally, this study included the traits of inflorescences, which have not been compared between three chemotypes in the literature. Phenotype markers identified in this study can facilitate preliminary cultivar identification and selection on live plants before or as a supplement to chemical and genetic analysis.**

Cannabis is an annual, normally dioecious, flowering plant with staminate plants tending to be taller than pistillate plants (Upton et al., 2014). The height can vary from 0.2 m to 6 m, with most of the plants reaching between 1 m to 3 m. Plant stems are erect, hollow, and grooved. *Cannabis* plant is taprooted, but taproots do not develop on vegetatively propagated plants. *Cannabis* plant has palmate leaves with 3 to 9 linear-lanceolate leaflets with serrations. Each female flower has one ovary encapsulated by bract and bracteoles (alternately called a *calyx*), out of which projects two long stigmas. Bract, bracteoles, stigmas, and small leaves that grow out of inflorescences leaves are densely covered by capitate stalked trichomes, where most cannabinoids and terpenoids are biosynthesized and stored. Trichomes are denser in female plants and decrease in density from inflorescences, leaves, stems, to roots (which are devoid of cannabinoids). *Cannabis* can host disease-causing pathogens, including *Botrytis cinerea*, which causes gray mold in the inflorescences (Clarke and Merlin, 2016; Kusari et al., 2013; McPartland, 1991, 1996; Punja et al., 2019).

Whether genus *Cannabis* is monotypic or polytypic is still a debate. Although several putative *Cannabis* species were discovered and proposed for distinguishing *Cannabis*, only two were widely accepted: *C. sativa* and *C. indica*. Linnaeus described *C. sativa* L. in *Species Plantarum* (Linnaeus, 1753) as having loose inflorescences covered with sparse trichomes and resembling a northern European fiber-type landrace (McPartland, 2017). Later in 1785, de Lamarck described a second (or sub-) species, *C. indica* Lam., which was collected in India, with dense trichomes, narrower leaflets, branching habitus, poorer fiber quality, harder stem, and thinner cortex, but stronger psychoactive effects (de Lamarck, 1785). Schultes traveled to Afghanistan in 1971 and described *C. indica* as having broad leaflets, densely branched with very dense inflorescences for hashish (resin) production, which deviated from Lamarck's original taxonomic concept (Schultes et al., 1974). Anderson drew illustrations of *C. indica* and *C. sativa*, the former represented as short, conical, densely branched, with broad leaflets and the latter as relatively tall, laxly branched, with narrow leaflets (Anderson, 1980), which aligned with Schultes but differed from Lamarck. *Cannabis* can be assigned as one of three chemotypes based on THC and CBD content (Small et al., 1973). Chemotype I is THC dominant, with more than 0.3% THC and less than 0.5% CBD. Chemotype II is intermediate, with high contents of both CBD (more than 0.5% THC) and THC (more than 0.3% THC). Chemotype III is CBD dominant with less than 0.3% THC. This quantitative approach was further developed into a qualitative measure using THC/CBD ratios: chemotype I has THC/CBD > 1, chemotype II has THC ≈ CBD, and chemotype III has THC/CBD < 1 (de Meijer et al., 2003; Mandolino et al., 2003; Turner et al., 1979).

Hillig carried out genetic, chemical, and morphological analysis on 157 accessions of diverse geographic origin before large-scale hybridization, classifying them into two species, *C. sativa* and *C. indica*, and seven putative taxa, including *C. indica* narrow-leaflet drug (NLD) biotype, *C. indica* wide-leaflet drug (WLD) biotype, *C. indica* hemp biotype, *C. indica* feral biotype, *C. sativa* hemp biotype, *C. sativa* feral biotype, and putative ruderal populations (Hillig, 2004, 2005a, 2005b, 2005c; Hillig and Mahlberg, 2004). The NLD biotype included landraces of Indian heritage (including varieties of the Indian subcontinent, Africa, and other drug-producing regions), corresponding to Lamarck's *C. indica*. The WLD biotype included landraces from Afghanistan and Pakistan, corresponding with Schultes's *C. indica*. The *C. indica* hemp biotype included landraces from southern and eastern Asia, while the *C. sativa* hemp biotype included landraces from Europe, Asia Minor, and Central Asia.

By the end of 1980, nearly all drug-type cannabis cultivated in the United States,

Canada, and Europe are cross-bred to achieve high THC content cultivars, called "sinsemilla" (meaning seedless) (McPartland, 2017). Cannabis breeders and users use vernacular "Sativa" to describe cultivars with narrow leaflets and "Indica" for cultivars with broad or wide leaflets, based on illustrations of Anderson, which deviated from the original botanical nomenclature (McPartland, 2017). Even so, researchers have tried to differentiate these two categories genetically and chemically (Lynch et al., 2016; Sawler et al., 2015; Vergara et al., 2016). However, these vernacular categories are unreliable for medical applications due to extensive cross-breeding and unreliable labeling during unrecorded hybridization (McPartland, 2017). CBD dominant and intermediate varieties were also excluded from these studies despite getting increasing attention from the therapeutic potential of CBD (Avraham et al., 2011; Bloomfield et al., 2020; French et al., 2017; McGuire et al., 2018), especially the indication of regulatory-approved prescription CBD (marketed as Epidiolex) to treat epilepsy (Billakota et al., 2019; Kaplan et al., 2017; Szaflarski et al., 2018). In addition to delimiting "Sativa" and "Indica" plants, recent studies tried to differentiate three chemotypes by profiling secondary metabolites (Jin et al., 2020), developing genetic markers (Borna et al., 2017; Kojoma et al., 2006; Pacifico et al., 2006; Rotherham et al., 2011; Toth et al., 2020), and comparing sequence and copy number variation of THC acid synthase and CBD acid synthase (McKernan et al., 2015; Onofri et al., 2015).

Although phenotypic differences are essential for delimiting plant species, this aspect of the plants has been largely ignored or limited when studying modern *Cannabis* cultivars. The description of leaflet shapes, if any, were determined visually and subjectively from sample providers without quantitative measures. There are no tangible data on determining what leaves are narrow leaflet and what leaves are wide leaflet, or whether an intermediate category exists between them due to hybridization. Lastly, plant morphology includes both qualitative and quantitative traits. Qualitative traits are ratios of two measurements, for example, the ratio of the width and length of a central leaflet on a node. Quantitative traits are absolute measurements, such as plant height and yield. Qualitative traits are usually determined by a single genetic locus, while quantitative traits usually result from interactions between several genes and environmental variables (Hillig, 2005b; McPartland, 2017). To study the phenotypic variation on plants, it is necessary to grow them in a single location under identical conditions to control environmental variation (Small et al., 1976).

The objectives of this study were the following:

1. Investigate whether modern *Cannabis* cultivars (including CBD dominant, intermediate, and THC dominant cultivars) can be differentiated using morphological traits.
2. Identify qualitative phenotypic markers, supplied with quantitative markers, that can be leveraged to select and distinguish chemotypes.

Materials and Methods

Plant material. In this project, 21 commercially available cultivars were grown in a commercial greenhouse (Fig. 1) under a research license issued by Health Canada. Where possible, the reported ancestry ("Sativa," "Indica," or "Sativa-dominant" and "Indica-dominant") was obtained from the Leafly online database (https://www.leafly.ca/) or from the licensed producer providing the cultivar (Table 1). Each cultivar was analyzed for chemical composition using methods established in previous work (Jin et al., 2020) and labeled as "THC dominant," "CBD dominant," or "intermediate." Five cuttings per cultivar were propagated in Jiffy 7 Peat Pellets Seed Starting Plugs (Jiffy, Pokemouche, NB, Canada) under SunBlaster T5HO fluorescent lighting (SunBlaster Holdings ULC, Langley, BC, Canada) and a 24-hour light photoperiod for 2 weeks. Eighty-fivesuccessful rooted plants were then transplanted to 6-inch pots with Dark Matter Super Soil (Destiny Grow Systems, Grand Forks, BC, Canada) for vegetative growth. Each plant was

Fig. 1. *Cannabis* plants grown in a greenhouse from rooting to flowering. (**A**) Five clones per cultivar rooted in Jiffy 7 plugs. (**B**) Vegetative traits were measured 40 d after rooting. (**C**) After 2 months of vegetative growth, plants were transplanted into planters with 12 plants per planter. (**D**) Light regime changed to 12 h per day and flowers began to grow. (**E**) After 2 months of flowering, plants were ready for harvest.

Received for publication 7 Dec. 2020. Accepted for publication 25 Jan. 2021.
Published online 25 February 2021.
We are grateful to PBG BioPharm Inc. for providing financial support for travel expenses and funding for publication.
D.J. and J.S. were employed by the company PBG BioPharma Inc.; P.H. was employed by the company Egret Bioscience and Lighthouse Genomics. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
D.J. conceived the project, designed the experiments, performed the experiments, collected and analyzed the data, and wrote the manuscript. P.H. found the resources of Cannabis cultivars and a commercial greenhouse for this project and proofread the manuscript. J.S. provided funding, provided suggestions, and proofread the manuscript. J.C. was the supervisory author and monitored the research progress, provided suggestions, and finalized the manuscript.
J.C. is the corresponding author. E-mail: jc65@ualberta.ca.

Table 1. Information and assignment of 21 cultivars into three chemotypes based on THC and CBD ratio.

Cultivar number	Cultivar name	Chemotypes	Total THC[z] ratio (mean ± SD)	Total CBD[z] ratio (mean ± SD)	Clusters	"Sativa" or "Indica"
1	Lemon Garlic OG	1-Intermediate	29.0% ± 2.4%	53.5% ± 2.2%	C1	"Indica" dominant
2	Royal Medic	2-Intermediate	32.7% ± 1.4%	50.9% ± 1.0%	C3	"Sativa" dominant
3	Blue Hawaiian	3-CBD	3.4% ± 0.2%	77.7% ± 0.8%	C3	"Sativa" dominant
4	Kandy Kush	4-CBD	3.9% ± 0.4%	77.5% ± 1.0%	C3	"Sativa" dominant
5	Special	5-CBD	3.3% ± 0.3%	78.6% ± 0.9%	C3	Not provided
6	NN	6-CBD	3.3% ± 0.2%	77.2% ± 0.5%	C3	Not provided
8	Treat	8-CBD	3.1% ± 0.5%	78.4% ± 1.4%	C3	Not provided
9	High	9-Intermediate	25.2% ± 1.1%	55.3% ± 0.7%	C3	Not provided
10	CB7	10-CBD	3.3% ± 0.7%	79.8% ± 1.2%	C3	Not provided
11	33°	11-THC	79.9% ± 1.0%	0.4% ± 0.2%	C1	Not provided
12	Banana Cake	12-THC	81.7% ± 0.4%	0.4% ± 0.1%	C2	"Indica" dominant
13	Bananium	13-THC	81.7% ± 0.6%	0.3% ± 0.05%	C3	"Indica" dominant
14	Burmese Blueberry	14-THC	78.8% ± 1.1%	0.3% ± 0.02%	C2	"Indica" dominant
15	Divine Banana	15-THC	81.7% ± 1.1%	0.3% ± 0.05%	C2	"Indica" dominant
16	Granddaddy Purple	16-THC	74.1% ± 0.7%	0.4% ± 0.1%	C2	"Indica" dominant
18	Lemon Sorbet	18-THC	84.4% ± 0.6%	0.6% ± 0.3%	C1	"Indica" dominant
19	MeatHead	19-THC	82.0% ± 1.2%	0.2% ± 0.03%	C2	"Indica" dominant
20	Nanitro	20-THC	78.6% ± 0.4%	0.3% ± 0.1%	C1	"Indica" dominant
21	Platinum Jelly Punch	21-THC	80.0% ± 1.0%	0.4% ± 0.1%	C1	"Indica" dominant
22	SBSK2 (Lemon Thai)	22-THC	79.7% ± 0.3%	0.4% ± 0.1%	C3	50/50 hybrid
23	Super Sherbet	23-THC	79.6% ± 1.6%	0.2% ± 0.02%	C1	"Indica" dominant

[z]Total THC = THCA × 0.877 + THC. Total CBD = CBDA × 0.877 + CBD.

Table 2. Phenotypic characteristics evaluated on each plant assigned to three chemotypes.

	Code	Characteristic	Unit/Notes
1	HgtVeg	Plant height 40 d after rooting	cm
2	DiaVeg	Stem diameter at base 40 d after rooting	mm
3	StmClrVeg	Reddish-brown coloration at base of stem of plants 40 d after rooting	Visually rated: 1, absent; 2, somewhat apparent; 3, present
4	VisGrnVeg	Visual determination of greenness 40 d after rooting	Visually rated: 1, light-green; 2, green; 3, deep-green
5	BranchVeg	Extent of branching 40 d after rooting	Visually rated: 1, less branching; 2, branching; 3, heavily branching
6	StretchVeg	Extent of stretching 40 d after rooting	Visually determined: 1, compact; 2, normal; 3, very stretching
7	NodeVeg	Number of nodes 40 d after rooting	
8	IntLngVeg	Mean internode length 40 d after rooting	mm
9	LftNumVeg	Mean leaflet number at node n 40 d after rooting, n = 3 to 10 (or highest number < 10)	
10	CtrLftLngVeg	Mean length of central leaflet at node n 40 d after rooting, n = 3 to 10 (or highest number < 10)	mm
11	CtrLftWdtVeg	Mean width of central leaflet at node n 40 d after rooting, n = 3 to 10 (or highest number < 10)	mm
12	LftRatioVeg	Mean width/length ratio of central leaflet at node n 40 d after rooting, n = 3 to 10 (or highest number < 10)	
13	LftShapeVeg	Mean ratio of distance from base of central leaflet to widest point/total length at node n, 40 d after rooting, n = 3 to 10 (or highest number < 10)	
14	PetLngVeg	Mean petiole length at node n 40 d after rooting, n = 3 to 10 (or highest number < 10)	mm
15	PetWdtVeg	Mean petiole width at node n 40 d after rooting, n = 3 to 10 (or highest number < 10)	mm
16	PetRatioVeg	Mean petiole width/thickness ratio at node n 40 d after rooting, n = 3 to 10 (or highest number < 10)	
17	PriSerVeg	Mean number of primary serrations on central leaflet at node n 40 d after rooting, n = 3 to 10 (or highest number < 10)	
18	SecSerVeg	Mean number of secondary serrations on central leaflet at node n 40 d after rooting, n = 3 to 10 (or highest number < 10)	
19	ChlphlVeg	Mean leaf chlorophyll concentration at node n 40 d after rooting, n = 3 to 10 (or highest number < 10)	
20	HgtFlw	Final height at the end of flowering stage	cm
21	HgtRat	Ratio of height 40 d after rooting over height at the end of flowering stage	cm
22	DiaFlw	Stem diameter at base at the end of the flowering stage	mm
23	StmClrFlw	Reddish-brown coloration at base of stem of plants at the end of the flowering stage	Visually rated: 1 = absent, 2 = somewhat apparent, 3 = present
24	YieldFlw	Flower yield per plant	g
25	WeightFlw	Mean weight per inflorescence averaged from at least ten inflorescences at the end of flowering stage	mg
26	OvrAprFlw	Overall appearance of inflorescences at the end of flowering stage	Visually rated: loose = 1, intermediate = 2, compact = 3
27	SugrLftClrFlw	Color of leaves in the inflorescence at the end of flowering stage	Visually rated: green = 1, mix of green and purple = 2, purple = 3
28	CalyxClrFlw	Color of calyx	Visually rated: green = 1, mix of green and purple = 2, purple = 3
29	ResinFlw	Whether inflorescences, on average of 5 from one plant, are resinous: sparkly, dense, sticky trichomes	Visually rated: 1. non-resin production, 2. intermediate, 3. resin production
30	SickFlw	Sickness at the end of flowering stage	Visually rated: *Botrytis cinerea* present = 1, absent = 0

numbered, then placed adjacent to one another in grids, the order of which was determined using a random-number generator. The random placement of plants was intentional to smooth out the impact of environmental variations within the room on the resultant data. Natural light was supplemented with artificial lighting using adjustable Gavita Pro 1000e DE HPS (Gavita, Vancouver, WA). The photosynthetically active radiation (*PAR*) was measured using an Apogee MQ-200 Quantum Separate Sensor (Apogee Instrument, Logan, UT) and was determined to be 200 ± 68 $\mu mol \cdot m^{-2} \cdot s^{-1}$ at the canopy level, on average. The greenhouse temperature was set constant at 22 °C. The plants were watered with Alaska Fish Fertilizer (N–P–K Ratio 5:1:1) (Alaska, Canada) every 3–4 d when the soil was visually dry. After 2 months' vegetative growth, plants were transferred to nine wooden planters, each measuring 150 cm × 150 cm and filled with soil, with 12 plants per planter and a 12-hour photoperiod to induce flowering. Natural light was supplemented with adjustable Gavita Pro 1000e DE HPS (Gavita). The average *PAR* at canopy level, measured over 3 d, was 559 ± 71 $\mu mol \cdot m^{-2} \cdot s^{-1}$ in the morning and 1159 ± 198 $\mu mol \cdot m^{-2} \cdot s^{-1}$ at noon. The highest *PAR* readings ranged from 1016 ± 295 to 1390 ± 104 $\mu mol \cdot m^{-2} \cdot s^{-1}$ at the canopy level for the nine planters, and the relative standard deviation was 11.9%, indicating a relatively even light distribution. The greenhouse temperature was set constant at 28 °C during this phase. Relative humidity ranged between 35% and 60%. The plants were watered every 3–4 d when the soil was visually dry. After 2 months of flowering, the whole plants were harvested and hung to dry in a closed environment. Horticultural fans were used to maintain air circulation, and the temperature was kept under 35 °C. The plants were dried for 7 d until the leaves and stems became brittle. At this time, the plants' moisture content is usually below 10% to 15% (mg/mg) (Caplan, 2018; Potter, 2009). Dried material was stored at room temperature until analysis.

Morphological traits evaluated. Traits on live plants were measured using a ruler and a digital micrometer 40 d after rooting and again at the end of the flowering phase (Table 2). Because it was difficult to evaluate leaf traits with respect to nodal positions in cuttings, mean values of traits on leaves were averaged from node 3 to node 10 (or highest node number if fewer than 10 nodes). Leaf traits were measured as illustrated in Fig. 2, referencing Anderson's method (Anderson, 1980; Hillig, 2005c). Petiole width and thickness were measured midway along the petiole. Leaf chlorophyll concentrations were measured using a MC-100 chlorophyll concentration meter (Apogee). After harvest and drying, the following traits were measured: inflorescence yield, mean inflorescence weight averaged from at least ten inflorescences, compactness of harvested inflorescences, whether the trichomes were dense and resinous, and incidence of *Botrytis cinerea* in inflorescences. All traits in Table 2 were treated as continuously variable for statistical analysis.

Statistical analysis. In total, 85 plants representing 21 cultivars were included in the following analysis. First, correlations were calculated between individual morphological traits with THC and CBD (ratios relatively to total cannabinoids), respectively. Ratios were employed because the relative compound proportions are more stable than absolute values, the latter changing between growth stages, plant parts, and environmental factors (Hillig, 2005c). Then, unsupervised (no preassigned categories as constraints) hierarchical clustering using Ward's minimum variance method (Ward, 1963) and principal component analysis (PCA) (Jolliffe, 2002) were used to check within-cultivar variation and between-cluster variation. Finally, the data were subjected to supervised canonical correlation analysis with preassigned chemotypes (Table 1). Canonical correlation analysis is a multiple discriminant analysis that calculates the correlation between preassigned clusters and the set of covariates describing the observations (morphological traits in this study) (Hotelling, 1936). Canonical variables are linear combination of the covariates that maximize the multiple correlation between the clusters, and the covariates are uncorrelated with each other. The analysis outputs a biplot with the first two canonical variables that provide maximum separation among the clusters. To identify phenotypic markers that contribute most to each chemotype, one-way analysis of variance (ANOVA) followed by Tukey's honestly significant difference (HSD) post hoc test at the 0.05 significance level were used to determine whether significant differences existed between all clusters and each pair of clusters. Statistical analysis was performed with JMP 14.0.0 (SAS, Cary, NC).

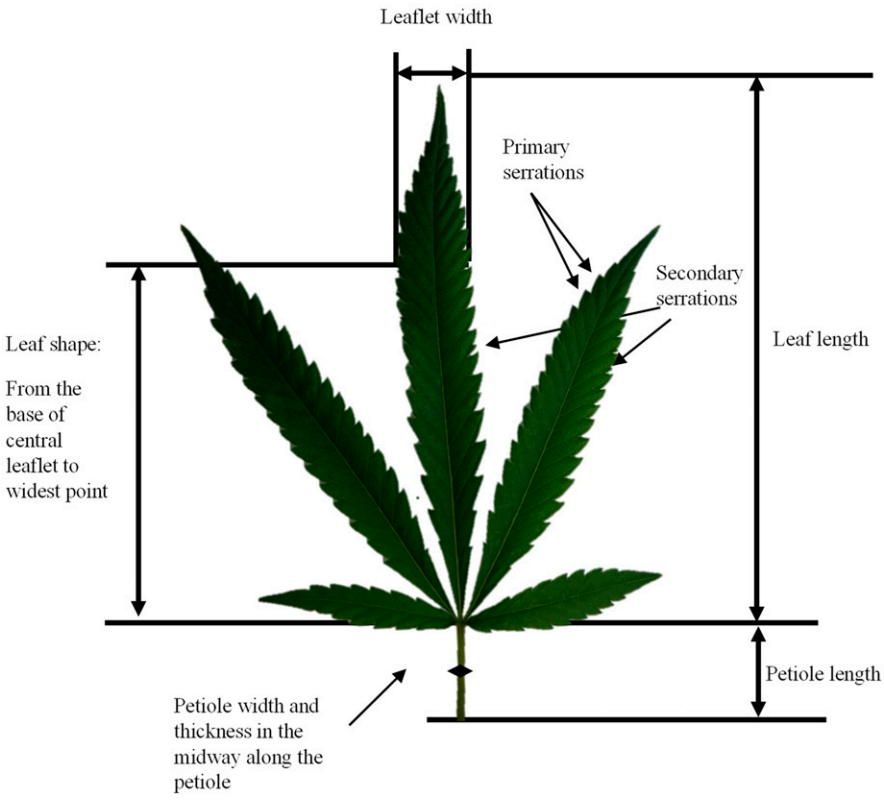

Fig. 2. Leaf traits measured on the central leaflet of fully expanded leaves while still attached to the living plant.

Results

Correlation analysis of morphological traits with THC and CBD. Correlations of THC and CBD with morphological traits are plotted in Fig. 3. The ratio of THC is positively correlated with leaf width/length ratio (0.77), width of central leaflet (0.55), final height at the end of flowering stage (0.54), inflorescence yield per plant (0.46), leaf chlorophyll concentration (0.42), sickness (*Botrytis cinerea* incidence) (0.41), stem diameter at base at the end of the flowering stage (0.33), and weight per inflorescence (0.32) in decreasing correlations. The ratio of CBD is positively correlated with the number of primary serrations on the central leaflet (0.52), number of leaflets (0.44), and length of central leaflet averaged from each node (0.35). The traits that were positively correlated with THC were all negatively correlated

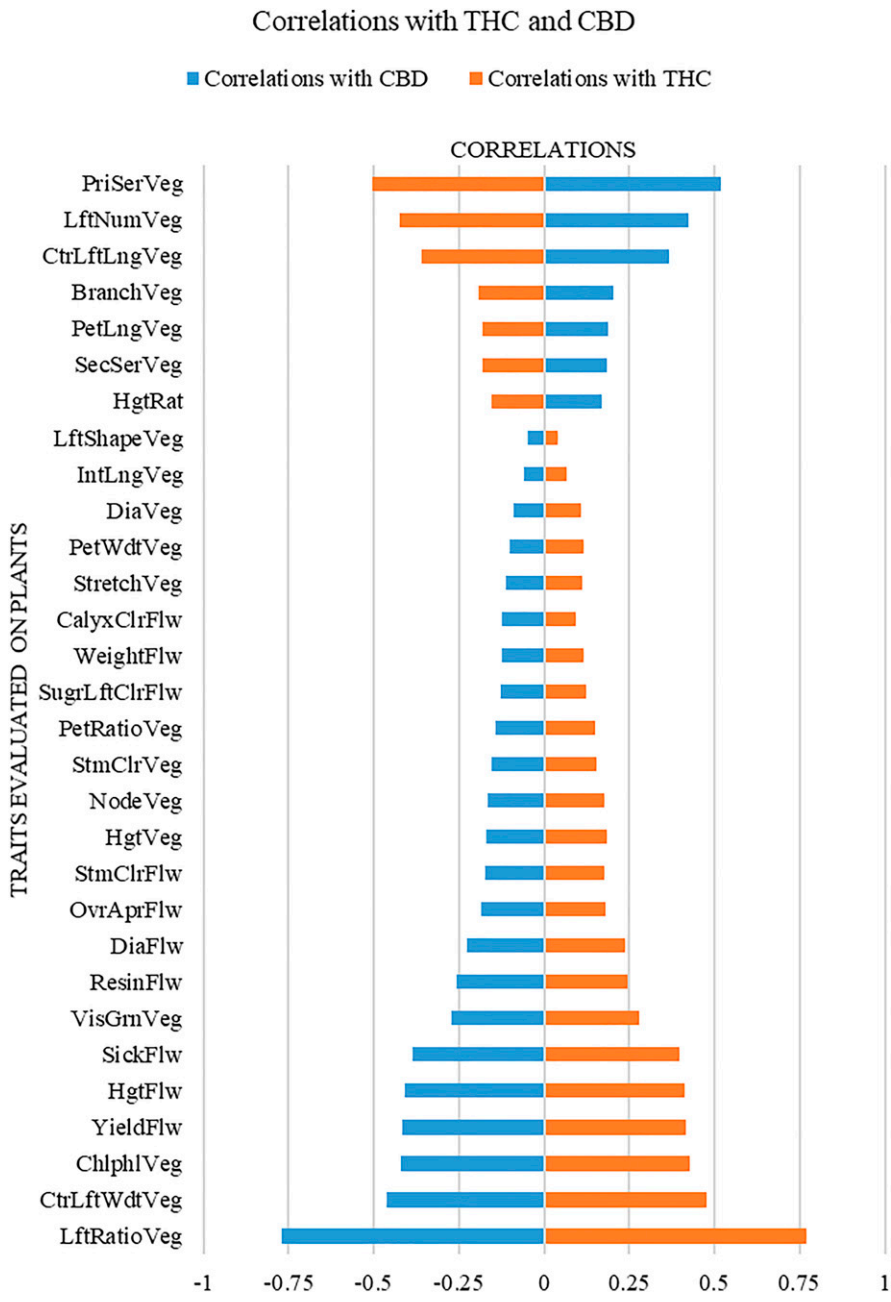

Fig. 3. Correlations of morphological traits with THC and CBD.

with CBD, where the absolute values of the correlations were similar, and vice versa. Interestingly, the correlations of the color of the calyx in inflorescences (green or purple) with CBD and THC are –0.12 and 0.09, respectively. The correlations of the color of the leaves in inflorescences (green or purple) with CBD and THC are –0.13 and 0.12, respectively. Neither was highly correlated with CBD and THC production.

Unsupervised hierarchical clustering. A hierarchical clustering dendrogram of the 85 plants is shown in Fig. 4, where most plants of the same cultivars were clustered together. Plants in 2-intermediate cultivar were scattered over the dendrogram. They seemed to experience different growth conditions, possibly related to uneven light interception. The dendrogram shows two major branches: one branch with plants from CBD dominant cluster C1 and the other with plants from THC dominant cluster C3. Plants from intermediate cluster C2 were distributed over both branches, showing phenotypic similarities with both THC dominant cultivars and CBD dominant cultivars, possibly due to hybridization.

Unsupervised principal component analysis. Figure 5 shows a scatterplot of 85 plants on PC1 and PC2. Plants of the same cultivars tended to occupy the same region on the plot, which shows small, within-cultivar variation and relative consistent morphological profiles within each cultivar. PC1 and PC2 explained 21.7% and 13.6% of total variance, respectively. These numbers are comparable to those of Hillig's study, where the numbers were 29.0% and 17.3%, respectively (Hillig, 2005c). Plants from THC dominant cluster C3 mainly occupied the right side the plot, while plants from CBD dominant cluster C1 occupied the left. Plants from intermediate cluster C2 occupied the middle of the plot and were mixed with both THC dominant plants and CBD dominant plants. Although plants from the same chemotypes tend to cluster together, the three clusters overlap, which may explain why PC1 and PC2 only explained 35.3% of the total variance. Cultivars assigned to C3 expressed a greater range of phenotypic variation than those assigned to C1 and C2 in the PC scatter plot, which may be due to a long history of selection for high THC levels for recreational purposes (McPartland, 2017).

The loading matrix in Table 3 lists the traits that contributed most to the separations along PC1 and PC2. Loadings with absolute values equal to or greater than 0.4 are listed in the table. PC1 was positively correlated with the width of central leaflet, final height at the end of flowering stage, leaflet width/length ratio, extent of stretching, inflorescences yield per plant, and so on, which were traits identified as positively correlated with THC. PC1 was negatively correlated with the number of primary and secondary serrations on central leaflet, which were traits identified as positively correlated with CBD. Traits that were positively correlated with PC2 and nearly vertical with PC1 included length of the central leaflet and the height ratio between 40 d after rooting and at the end of flowering stage. These traits overlap with plants from intermediate C2 in the scatter plot. Traits positively correlated with PC2 and positively correlated with PC1 were internode length, plant height 40 d after rooting, and petiole width. These traits are responsible for the location of the THC dominant plants on the upper right quadrant. The number of primary serrations on the central leaflet was positively correlated with PC2 and negatively with PC1, which was responsible for CBD dominant plants located on the upper left quadrant of the plot.

Supervised canonical correlation analysis. The canonical correlation analysis showed good separation between the preassigned chemotypes (Fig. 6). Each plant was predicted to be in its originally preassigned clusters C1, C2, and C3 with a 92.9% (79/85) accuracy (Table 4). Means, standard deviations (±SD), ranges, Tukey HSD multiple tests at the 0.05 significance level, and *P* value of one-way ANOVA of 30 traits for each of the three clusters were calculated (Table 5). The largest number of significant differences was 19, which was between C1 and C3. The most similar pair was C1 and C2, with three significant differences. The number of significant differences between C2 and C3 was six. CBD dominant cultivars had more leaflets, longer central leaflets, and more primary and secondary serrations, which are traits

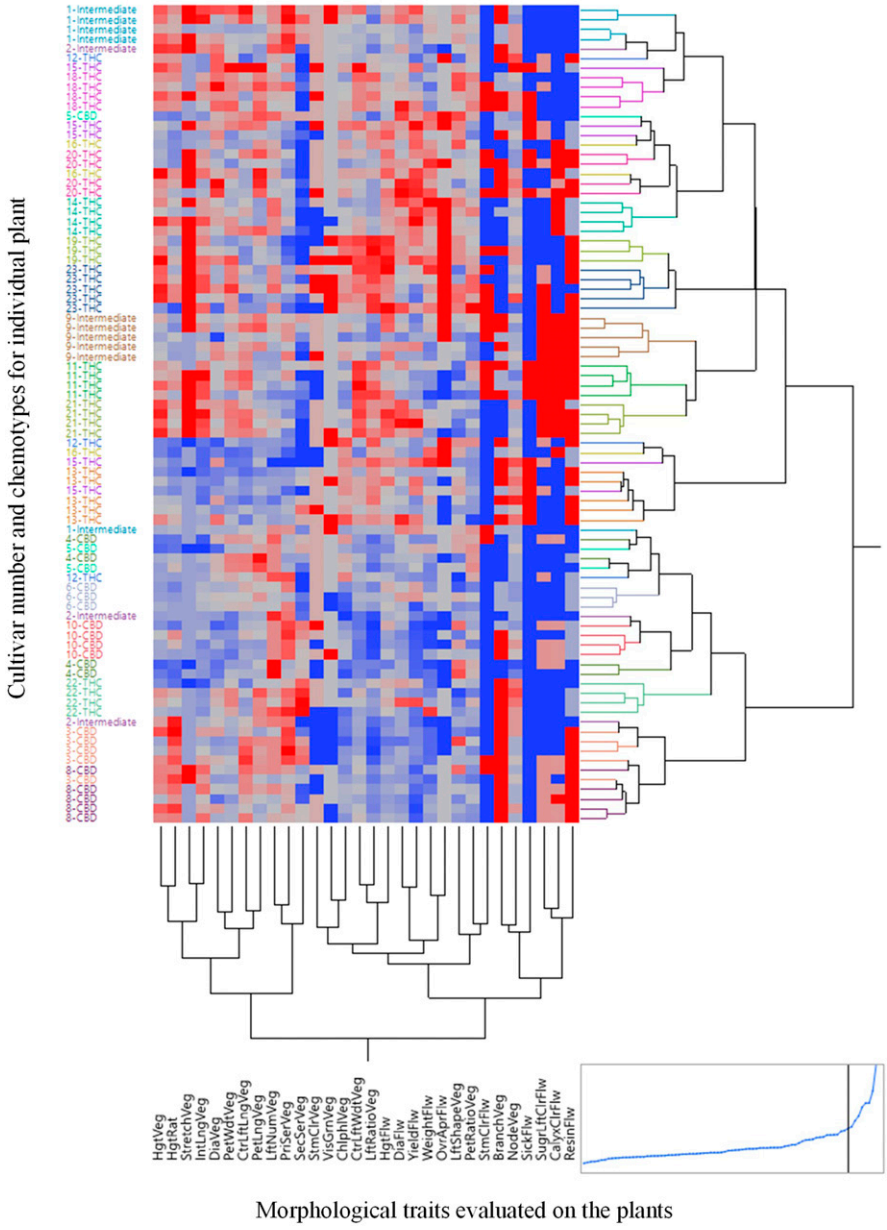

Fig. 4. Dendrogram by hierarchical clustering analysis using 30 morphological traits of 85 plants representing 21 cultivars.

positively correlated with CBD production. THC dominant cultivars had higher plant height, larger stem diameter, deeper-green color of leaves, more nodes, a wider central leaflet, larger width/length ratio of the central leaflet, wider petiole width, higher chlorophyll concentrations, higher inflorescences yield, larger inflorescence weight, more compact-looking inflorescence, sticky inflorescences, and higher *Botrytis cinerea* incidence. These traits were also positively correlated with THC production. Most traits for intermediate cultivars were at an intermediate level between C1 and C3 or at the same level with cultivars in C1 or C3. Several traits were not significantly different between three groups, including reddish-brown coloration at the base of the stem, extent of branching, internode length, leaf shape, petiole length, petiole width/thickness ratio, plant growth rate, color of leaves, and color of calyx. Example leaflets for C1, C2, and C3, green/purple leaves and calyx, inflorescences infected with *Botrytis cinerea* are shown in Fig. 7.

Discussion

In this study, 21 *Cannabis* cultivars belonging to three chemotypes (THC dominant, intermediate, and CBD dominant) were grown in a greenhouse. Morphological traits were measured, and canonical correlation analysis was used to test the goodness of fit between chemotype labeling and phenotypic variations. This study also identified phenotypic markers for each chemotype. The widespread crossbreeding and introgression in *Cannabis* blurred the differences between NLD and WLD cultivars, as well as made the distinction between their hybrids difficult (McPartland, 2017). However, useful suites of traits were identified for differentiating three chemotypes. These identified traits were largely consistent as confirmed by correlation analysis, PCA, and canonical correlation analysis. The ratio-based qualitative differences may be more consistent between growing environments and therefore more useful for differentiation when applied to conditions different from those used in this study. THC dominant cultivars had the largest mean width/length ratio of 0.25 ± 0.03. Hillig described two drug types (THC dominant) before large-scale hybridization in the 1990s: NLD with width/length ratio ranged from 0.15 ± 0.02 to 0.24 ± 0.03 (measured with respect to nodal positions), and WLD with width/length ratio ranged from 0.22 ± 0.03 to 0.39 ± 0.06 (Hillig, 2005c). The THC dominant cultivars described in this study had width/length ratios at the high end of the ratio for NLD cultivars and the lower end of the ratio for WLD cultivars. This may be attributed to hybridization between NLD and WLD cultivars, which was performed to obtain sinsemilla hybrids with high THC content, low CBD content, and high inflorescences yield (Clarke and Merlin, 2016). This study is the first to describe leaf traits for CBD dominant and intermediate cultivars. CBD dominant cultivars had narrow leaflets with width/length ratios of 0.18 ± 0.02, whereas intermediate cultivars had intermediate width/length ratios of 0.20 ± 0.02. A complete genome assembly of CBD dominant cultivars revealed that these cultivars were created by integrating hemp-type CBD acid synthase gene clusters into a background of drug-type *Cannabis* to elevate CBD production (Kovalchuk et al., 2020). The intermediate width/length ratio for intermediate cultivars was likely a result of hybridization between purebred CBD dominant cultivars and THC dominant cultivars.

In addition to qualitative criteria, this study provided additional quantitative and visual phenotypic criteria that may have differentiation power. For example, CBD dominant cultivars were a lighter green, THC dominant cultivars were a deeper green, and intermediate cultivars had shades in between. Because morphological traits change depending on environment variables, absolute measurements, including the plant height, the leaflet length, and the inflorescence yield, may differ if plants are grown in a different environment. The visual greenness of leaf colors and chlorophyll concentrations were positively correlated with each other. The inflorescences of NLD/BLD hybrid cultivars in C3 were large and compact—a clear result of artificial selection. However, the major horticultural drawback of these hybrids is their susceptibility to fungal infections (Clarke and Merlin, 2016). NLD landraces originated

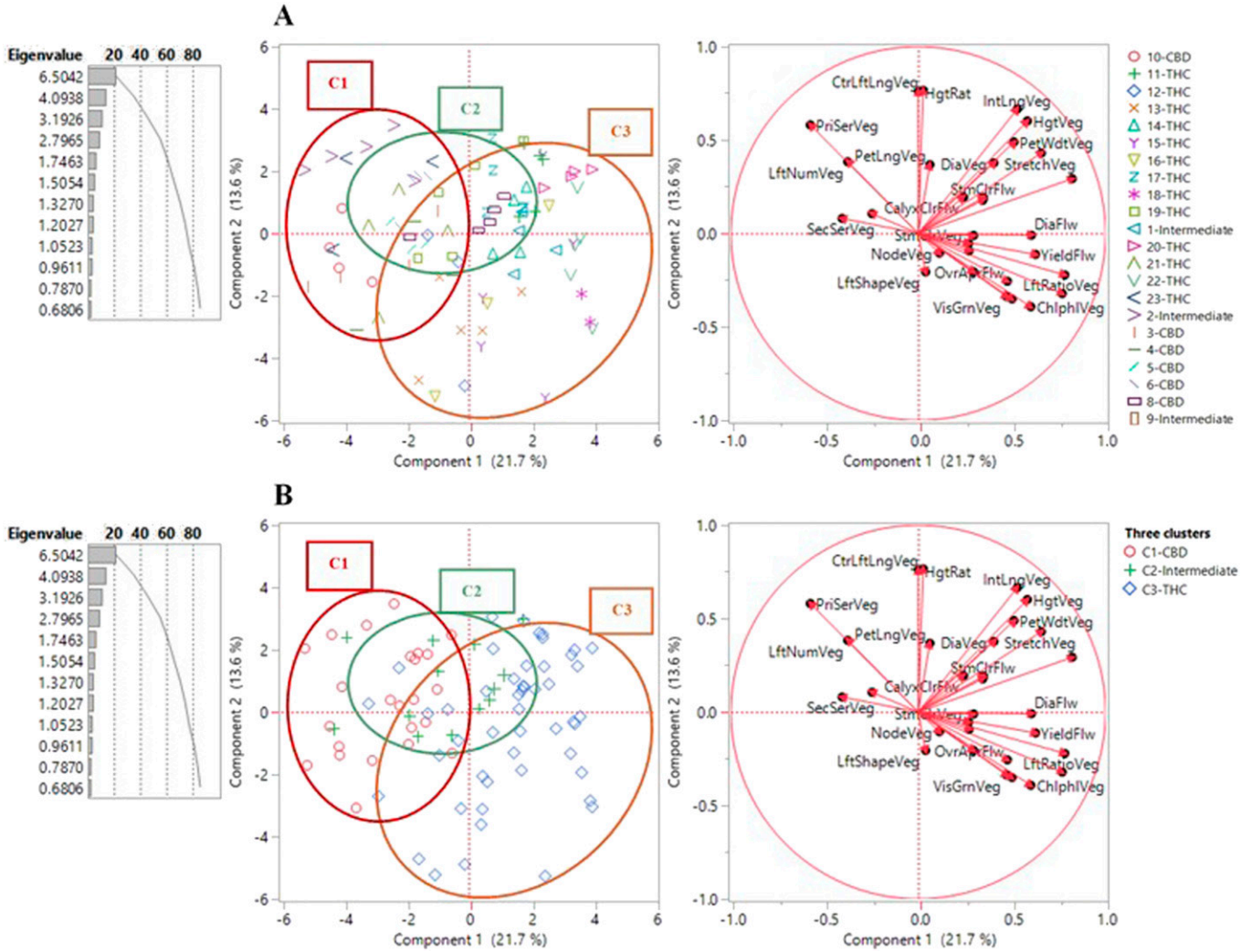

Fig. 5. PCA scatter plot (left) and loading plot (right) using 30 morphological traits of 85 plants representing 21 cultivars (**A**) using individual cultivar as label and (**B**) using three chemotypes as label.

Table 3. Formatted loading matrix for PC1 and PC2 (only compounds with absolute loadings > 0.4 are listed).

PC1				PC2			
Traits	Positive loadings	Traits	Negative loadings	Traits	Positive loadings	Traits	Negative loadings
CtrLftWdtVeg	0.82	PriSerVeg	–0.58	CtrLftLngVeg	0.77		
HgtFlw	0.78	SecSerVeg	–0.41	HgtRat	0.76		
LftRatioVeg	0.77			IntLngVeg	0.67		
StretchVeg	0.66			HgtVeg	0.60		
YieldFlw	0.63			PriSerVeg	0.58		
DiaFlw	0.60			PetWdtVeg	0.49		
ChlphlVeg	0.60			StretchVeg	0.43		
HgtVeg	0.58						
IntLngVeg	0.53						
PetWdtVeg	0.51						
WeightFlw	0.50						
OvrAprFlw	0.48						
VisGrnVeg	0.47						
DiaVeg	0.40						
BranchVeg	0.40						

from regions with relatively humid conditions (Colombia, India, Jamaica, Thailand, and so forth) and evolved natural resistance to fungal infection, whereas WLD landraces adapted to the arid Afghani environment in which fungal resistance was unnecessary (Clarke and Merlin, 2016; McPartland, 1996). More varieties in C3 had *Botrytis cinerea* infection than in C1 and C2, indicating that these cultivars face additional risk in a greenhouse environment. Alternatively, this may reflect the current market in

Table 4. CCA summary of preassigned and predicted classifications of 85 plants into three chemotypes using 30 traits.

Preassigned	Predicted		
Cluster A	C1-CBD	C2-Intermediate	C3-THC
C1-CBD	21	3	0
C2-Intermediate	1	12	0
C3-THC	0	2	46

North America: growers and users favor large and compact inflorescences, which are a trait of wide leaflet cultivars.

In summary, the identified suites of phenotypic signatures in this work can be used to determine chemotypes on live plants before or as a supplement to chemical and genetic analysis.

Table 5. Means (±SD) and ranges (mean-SD, mean+SD) of 30 traits for 85 plants assigned to C1-CBD dominant, C2-intermediate, and C3-THC dominant.[z]

		C1-CBD	C2-Intermediate	C3-THC	ANOVA
	Three clusters	N = 24	N = 13	N = 48	*P*
Vegetative growth (40 d after rooted)	HgtVeg (cm)	19.64 ± 5.57 b 14.07–25.21	23.47 ± 3.76 ab 19.71–27.23	24.05 ± 4.79 a 19.26–28.84	0.0012
	DiaVeg (mm)	3.26 ± 0.46 b 2.80–3.72	3.57 ± 0.63 ab 2.94–4.20	3.68 ± 0.61 a 3.07–4.29	0.013
	StmClrVeg	1.67 ± 0.56 a 1.11–2.23	2.00 ± 0.41 a 1.59–2.41	1.92 ± 0.64 a 1.28–2.56	0.1847
	VisGrnVeg	1.71 ± 0.55 b 1.16–2.26	2.15 ± 0.55 a 1.60–2.70	2.17 ± 0.47 a 1.70–2.64	0.0016
	BranchVeg	2.29 ± 0.85 a 1.44–3.14	2.46 ± 0.66 a 1.80–3.12	2.06 ± 0.91 a 1.15–2.97	0.35
	StretchVeg	2.00 ± 0.42 b 1.58–2.42	2.46 ± 0.52 a 1.92–2.98	2.44 ± 0.61 a 1.83–3.05	0.0036
	NodeVeg	8.04 ± 1.33 b 6.71–9.37	8.62 ± 1.12 ab 7.50–9.74	9.33 ± 1.77 a 7.56–11.1	0.008
	IntLngVeg (mm)	24.32 ± 5.10 a 19.22–29.42	27.46 ± 4.67 a 22.79–32.13	26.52 ± 6.58 a 19.94–33.10	0.1715
	LftNumVeg	4.92 ± 0.47 a 4.45–5.39	4.81 ± 0.52 ab 4.29–5.33	4.34 ± 0.74 b 3.60–5.08	0.0009
	CtrLftLngVeg (mm)	95.64 ± 10.93 a 84.71–106.57	95.77 ± 13.69 ab 82.08–109.45	84.74 ± 16.24 b 68.51–100.98	0.0071
	CtrLftWdtVeg (mm)	17.53 ± 2.39 b 15.14–19.92	19.31 ± 2.66 b 16.65–21.97	22.65 ± 3.51 a 19.14–26.16	<0.0001
	LftRatioVeg	0.18 ± 0.02 b 0.16–0.20	0.20 ± 0.02 b 0.18–0.22	0.25 ± 0.03 a 0.22–0.28	<0.0001
	LftShapeVeg	0.51 ± 0.04 a 0.47–0.55	0.51 ± 0.02 a 0.49–0.53	0.50 ± 0.04 a 0.46–0.54	0.9282
	PetLngVeg (mm)	40.20 ± 6.80 a 33.40–46.99	38.69 ± 6.82 a 31.87–45.51	38.93 ± 8.47 a 30.46–47.41	0.7756
	PetWdtVeg (mm)	1.34 ± 0.15 b 1.19–1.49	1.41 ± 0.16 ab 1.25–1.57	1.47 ± 0.21 a 1.26–1.68	0.0232
	PetRatioVeg	0.99 ± 0.03 a 0.96–1.02	1.00 ± 0.03 a 0.97–1.03	1.01 ± 0.04 a 0.97–1.05	0.2934
	PriSerVeg	15.04 ± 1.74 a 13.30–16.78	15.97 ± 1.49 a 14.48–17.46	12.75 ± 2.06 b 10.69–14.81	<0.0001
	SecSerVeg	1.14 ± 0.87 a 0.27–2.01	0.99 ± 1.04 a 0–2.03	0.34 ± 0.12 b 0.22–0.46	0.0002
	ChlphlVeg	16.92 ± 2.49 b 14.43–19.41	18.33 ± 3.00 ab 15.33–21.33	20.39 ± 3.38 a 17.01–23.77	<0.0001
At the end of flowering stage	HgtFlw (cm)	97.21 ± 15.00 b 82.21–112.21	115.54 ± 17.67 a 97.87–133.21	126.56 ± 13.42 a 113.14–139.98	<0.0001
	HgtRat	0.21 ± 0.06 a 0.15–0.27	0.21 ± 0.05 a 0.16–0.26	0.19 ± 0.04 a 0.15–0.23	0.5352
	DiaFlw (mm)	13.65 ± 2.41 b 11.24–16.06	15.58 ± 2.61 ab 12.97–18.19	16.85 ± 2.95 a 13.90–19.80	<0.0001
	StmClrFlw	1.13 ± 0.34 a 0.79–1.47	1.31 ± 0.48 a 0.83–1.79	1.31 ± 0.48 a 0.83–1.79	0.1509
	YieldFlw (g)	36.43 ± 26.73 b 9.70–63.16	60.40 ± 30.50 ab 29.90–90.90	85.09 ± 38.45 a 46.64–123.54	<0.0001
Inflorescences (Harvested)	WeightFlw (mg)	355.04 ± 130.40 b 224.64–485.44	411.69 ± 118.04 ab 293.65–529.73	542.75 ± 208.40 a 334.35–751.15	0.0002
	OvrAprFlw	1.71 ± 0.55 b 1.16–2.26	2.00 ± 0.71 ab 1.29–2.71	2.19 ± 0.74 a 1.45–2.93	0.0345
	SugrLftClrFlw	1.50 ± 0.51 a 0.99–2.01	1.77 ± 1.01 a 0.76–2.78	1.73 ± 0.83 a 0.90–2.56	0.4277
	CalyxClrFlw	1.50 ± 0.59 a 0.91–2.09	1.77 ± 1.01 a 0.76–2.78	1.77 ± 0.97 a 0.90–2.56	0.4341
	ResinFlw	2.17 ± 0.82 ab 1.35–2.99	1.77 ± 1.01 b 0.76–2.78	2.40 ± 0.72 a 1.68–3.12	0.0489
	SickFlw	0.04 ± 0.20 b 0–0.24	0.00 ± 0.00 b 0.00–0.00	0.40 ± 0.49 a 0–0.89	0.0002

[z]Means for the same trait not connected by same letter are significantly different.

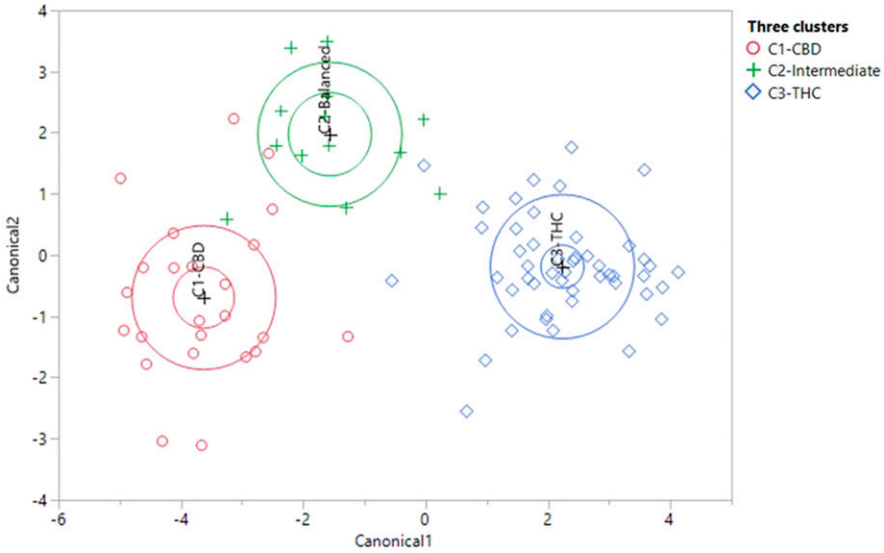

Fig. 6. Canonical correlation analysis scatterplot of 85 plants representing 21 cultivars on the first and second canonical axes using 30 traits. The cultivars were preassigned to three chemotypes in Table 1. The observations and the multivariate means of each group ("+") are represented as points on the biplot. A 95% confidence level ellipse is plotted for each mean. An ellipse denoting a 50% contour is plotted for each group, which contains ≈50% of the observations.

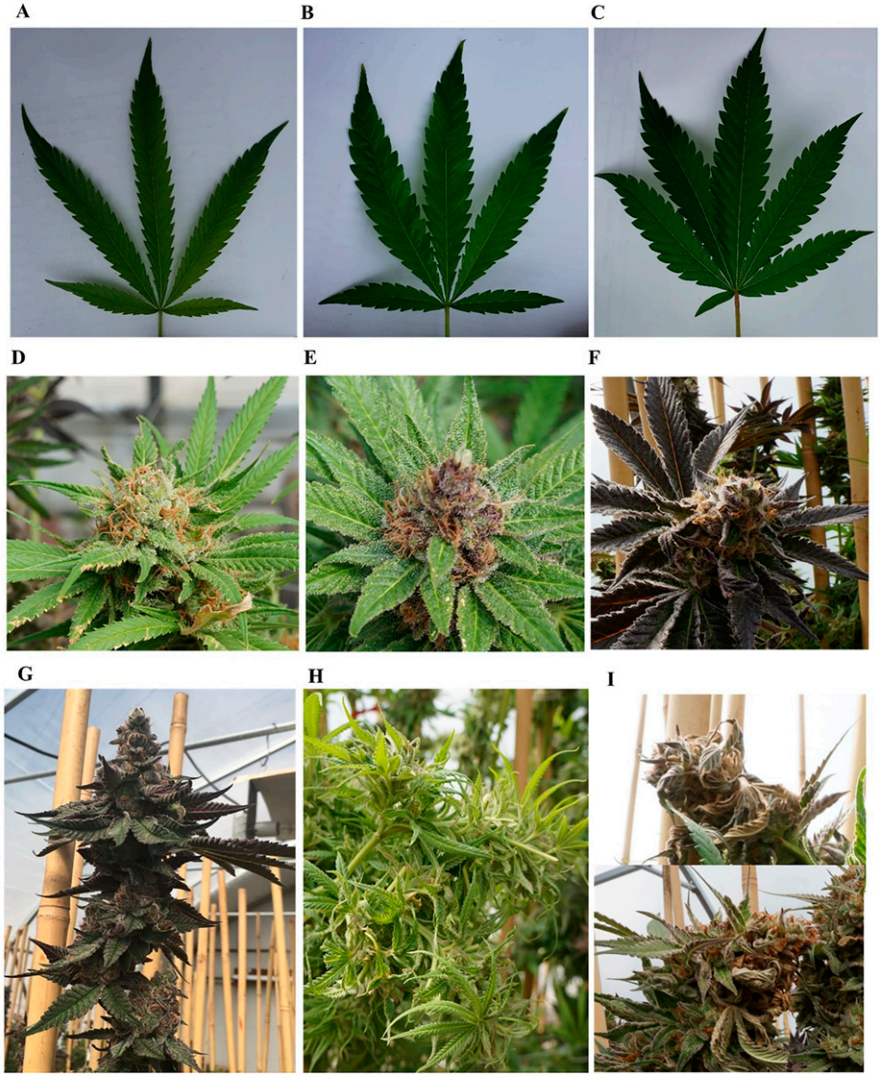

Fig. 7. (**A**) Example leaflets for CBD dominant cultivars, (**B**) intermediate cultivars, and (**C**) THC dominant cultivars. (**D**) Example mature inflorescences of cultivars with green leaves and green calyx, (**E**) green leaves with purple calyx, and (**F**) purple leaves with purple calyx. (**G**) Example of compact inflorescences, (**H**) loose inflorescences, and (**I**) inflorescences infected with *Botrytis cinerea*.

Literature Cited

Anderson, L.C. 1980. Leaf variation among *Cannabis* species from a controlled garden. Bot. Mus. Leaflf. Harv. Univ. 28(1):61–69.

Avraham, Y., N.C. Grigoriadis, T. Poutahidis, L. Vorobiev', I. Magen, Y. Ilan, R. Mechoulam, and E.M. Berry. 2011. Cannabidiol improves brain and liver function in a fulminant hepatic failure-induced model of hepatic encephalopathy in mice. Brit. J. Pharmacol. 162(7):1650–1658.

Billakota, S., O. Devinsky, and E. Marsh. 2019. Cannabinoid therapy in epilepsy. Curr. Opin. Neurol. 32(2):220–226, doi: 10.1097/WCO.0000000000000660.

Bloomfield, M.A.P., S.F. Green, C. Hindocha, Y. Yamamori, J.L.L. Yim, A.P.M. Jones, H.R. Walker, P. Tokarczuk, B. Statton, O.D. Howes, H.V. Curran, and T.P. Freeman. 2020. The effects of acute cannabidiol on cerebral blood flow and its relationship to memory: An arterial spin labelling magnetic resonance imaging study. J. Psychopharmacol. 34(9):981–989, doi: 10.1177/0269881120936419.

Borna, T., S.B. Tittes, S.A. Salami, and M. Shokrpour. 2017. High resolution melting curve analysis revealed SNPs in major cannabinoid genes associated with drug and non-drug types of cannabis. Biotechnol. Biotechnol. Equip. 31(4):839–845, doi: 10.1080/13102818.2017.1333456.

Caplan, D.M. 2018. Propagation and root zone management for controlled environment cannabis production (PhD Thesis). University of Guelph, Guelph, Ontario, Canada.

Clarke, R.C. and M.D. Merlin. 2016. Cannabis domestication, breeding history, present-day genetic diversity, and future prospects. Crit. Rev. Plant Sci. 35(5–6):293–327, doi: 10.1080/07352689.2016.1267498.

de Lamarck, J.B. 1785. Encyclopédie méthodique de botanique, vol. 1, pt. 2. Panckoucke, Paris, France, p. 694–695.

de Meijer, E.P.M., M. Bagatta, A. Carboni, P. Crucitti, V.M.C. Moliterni, P. Ranalli, and G. Mandolino. 2003. The inheritance of chemical phenotype in *Cannabis sativa* L. Genetics 163(1):335–346.

French, J., E. Thiele, M. Mazurkiewicz-Beldzinska, S. Benbadis, E. Marsh, C. Joshi, C. Roberts, A. Taylor, and K. Sommerville. 2017. Cannabidiol (CBD) significantly reduces drop seizure frequency in Lennox-Gastaut syndrome (LGS): Results of a multi-center, randomized, double-blind, placebo controlled trial (GWPCARE4) (S21.001). Neurology 88(16):S21–S001.

Hillig, K.W. 2004. A chemotaxonomic analysis of terpenoid variation in *Cannabis*. Biochem. Syst. Ecol. 32(10):875–891, doi: 10.1016/j.bse.2004.04.004.

Hillig, K.W. 2005a. Genetic evidence for speciation in *Cannabis* (Cannabaceae). Genet. Resources Crop Evol. 52(2):161–180.

Hillig, K.W. 2005b. A combined analysis of agronomic traits and allozyme allele frequencies for 69 *Cannabis* accessions. J. Ind. Hemp 10(1):17–30, doi: 10.1300/J237v10n01_03.

Hillig, K.W. 2005c. A systematic investigation of *Cannabis*, PhD thesis, Indiana University (PhD Thesis), Bloomington, IN.

Hillig, K.W. and P.G. Mahlberg. 2004. A chemotaxonomic analysis of cannabinoid variation in *Cannabis* (Cannabaceae). Amer. J. Bot. 91(6):966–975.

Hotelling, H. 1936. Relations between two sets of variates. Biometrika 28(3–4):321–377, doi: 10.1093/biomet/28.3-4.321.

Jin, D., K. Dai, Z. Xie, and J. Chen. 2020. Secondary metabolites profiled in cannabis inflorescences, leaves, stem barks, and roots for medicinal purposes. Sci. Rep. 10(1):3309, doi: 10.1038/s41598-020-60172-6.

Jolliffe, I.T. 2002. Principal component analysis. 2nd ed. Springer series in statistics. Springer-Verlag, New York, doi: 10.1007/b98835.
Kaplan, E.H., E.A. Offermann, J.W. Sievers, and A.M. Comi. 2017. Cannabidiol treatment for refractory seizures in Sturge–Weber syndrome. Pediatr. Neurol. 71:18–23.
Kojoma, M., H. Seki, S. Yoshida, and T. Muranaka. 2006. DNA polymorphisms in the tetrahydrocannabinolic acid (THCA) synthase gene in "drug-type" and "fiber-type" Cannabis sativa L. Forensic Sci. Int. 159(2):132–140, doi: 10.1016/j.forsciint.2005.07.005.
Kovalchuk, I., M. Pellino, P. Rigault, R. van Velzen, J. Ebersbach, J.R. Ashnest, M. Mau, M.E. Schranz, J. Alcorn, R.B. Laprairie, J.K. McKay, C. Burbridge, D. Schneider, D. Vergara, N.C. Kane, and T.F. Sharbel. 2020. The genomics of cannabis and its close relatives. Annu. Rev. Plant Biol. 71(1):713–739, doi: 10.1146/annurev-arplant-081519-040203.
Kusari, P., S. Kusari, M. Spiteller, and O. Kayser. 2013. Endophytic fungi harbored in *Cannabis sativa* L.: Diversity and potential as biocontrol agents against host plant-specific phytopathogens. Fungal Divers. 60(1):137–151, doi: 10.1007/s13225-012-0216-3.
Linnaeus, C. 1753. Species plantarum. 2:1057. In: Laurentii Salvii, Stockholm, Sweden.
Lynch, R.C., D. Vergara, S. Tittes, K. White, C.J. Schwartz, M.J. Gibbs, T.C. Ruthenburg, K. deCesare, D.P. Land, and N.C. Kane. 2016. Genomic and chemical diversity in *Cannabis*. Crit. Rev. Plant Sci. 35(5–6):349–363.
Mandolino, G., M. Bagatta, A. Carboni, P. Ranalli, and E. de Meijer. 2003. Qualitative and quantitative aspects of the inheritance of chemical phenotype in *Cannabis*. J. Ind. Hemp 8(2):51–72.
McGuire, P., P. Robson, W.J. Cubala, D. Vasile, P.D. Morrison, R. Barron, A. Taylor, and S. Wright. 2018. Cannabidiol (CBD) as an adjunctive therapy in schizophrenia: A multicenter randomized controlled trial. Amer. J. Psychiatry 175(3):225–231, doi: 10.1176/appi.ajp.2017.17030325.
McKernan, K.J., Y. Helbert, V. Tadigotla, S. McLaughlin, J. Spangler, L. Zhang, and D. Smith. 2015. Single molecule sequencing of THCA synthase reveals copy number variation in modern drug-type *Cannabis sativa* L. bioRxiv 028654, doi: 10.1101/028654.
McPartland, J.M. 2017. *Cannabis sativa* and *Cannabis indica* versus "Sativa" and "Indica," p. 101–121. In: S. Chandra, H. Lata, and M. A. ElSohly. *Cannabis sativa* L.—Botany and biotechnology. Springer, Chaum, Switzerland.
McPartland, J.M. 1996. McPartland: A review of *Cannabis* diseases. J. Intl. Hemp Assoc. 3(1):19–23.
McPartland, J.M. 1991. McPartland: Common names for diseases of *Cannabis sativa* L. Plant Dis. 75:226–227.
Onofri, C., E.P.M. de Meijer, and G. Mandolino. 2015. Sequence heterogeneity of cannabidiolic- and tetrahydrocannabinolic acid-synthase in *Cannabis sativa* L. and its relationship with chemical phenotype. Phytochemistry 116:57–68, doi: 10.1016/j.phytochem.2015.03.006.
Pacifico, D., F. Miselli, M. Micheler, A. Carboni, P. Ranalli, and G. Mandolino. 2006. Genetics and Marker-assisted Selection of the Chemotype in *Cannabis sativa* L. Mol. Breed. 17(3):257–268, doi: 10.1007/s11032-005-5681-x.
Potter, D.J. 2009. The propagation, characterisation and optimisation of *Cannabis sativa* L. as a phytopharmaceutical (PhD Thesis). King's College, London.
Punja, Z.K., D. Collyer, C. Scott, S. Lung, J. Holmes, and D. Sutton. 2019. Pathogens and molds affecting production and quality of *Cannabis sativa* L. Front. Plant Sci. 10:1120, doi: 10.3389/fpls.2019.01120.
Rotherham, D. and S.A. Harbison. 2011. Differentiation of drug and non-drug Cannabis using a single nucleotide polymorphism (SNP) assay. Forensic Sci. Intl. 207(1–3):193–197, doi: 10.1016/j.forsciint.2010.10.006.
Sawler, J., J.M. Stout, K.M. Gardner, D. Hudson, J. Vidmar, L. Butler, J.E. Page, and S. Myles. 2015. The genetic structure of marijuana and hemp. PLoS One 10(8):e0133292.
Schultes, R.E., W.M. Klein, T. Plowman, and T.E. Lockwood. 1974. *Cannabis*: An example of taxonomic neglect. Bot. Mus. Lealf. Harv. Univ. 23:337–367.
Small, E. and H.D. Beckstead. 1973. Letter: Cannabinoid phenotypes in *Cannabis sativa*. Nature 245(5421):147–148, doi: 10.1038/245147a0.
Small, E., P.Y. Jui, and L.P. Lefkovitch. 1976. A numerical taxonomic analysis of cannabis with special reference to species delimitation. Syst. Bot. 1(1):67–84, doi: 10.2307/2418840.
Szaflarski, J.P., E.M. Bebin, A.M. Comi, A.D. Patel, C. Joshi, D. Checketts, J.C. Beal, L.C. Laux, L.M. De Boer, M.H. Wong, M. Lopez, O. Devinsky, P.D. Lyons, P.P. Zentil, and R. Wechsler. CBD EAP study group. 2018. Long-term safety and treatment effects of cannabidiol in children and adults with treatment-resistant epilepsies: Expanded access program results. Epilepsia 59(8):1540–1548, doi: 10.1111/epi.14477.
Toth, J.A., G.M. Stack, A.R. Cala, C.H. Carlson, R.L. Wilk, J.L. Crawford, D.R. Viands, G. Philippe, C.D. Smart, J.K.C. Rose, and L.B. Smart. 2020. Development and validation of genetic markers for sex and cannabinoid chemotype in *Cannabis sativa* L. GCB Bioenergy 12(3):213–222, doi: 10.1111/gcbb.12667.
Turner, C.E., M.A. ElSohly, P.C. Cheng, and G. Lewis. 1979. Constituents of *Cannabis sativa* L., XIV: Intrinsic problems in classifying cannabis based on a single cannabinoid analysis. J. Nat. Prod. 42(3):317–319, doi: 10.1021/np50003a017.
Upton, R., L. Craker, M. ElSohly, A. Romm, E. Russo, and M. Sexton. 2014. Cannabis inflorescence: Cannabis spp.; Standards of identity, analysis, and quality control. American Herbal Pharmacopoeia, Scotts Valley, CA.
Vergara, D., H. Baker, K. Clancy, K.G. Keepers, J.P. Mendieta, C.S. Pauli, S.B. Tittes, K.H. White, and N.C. Kane. 2016. Genetic and genomic tools for *Cannabis sativa*. Crit. Rev. Plant Sci. 35(5–6):364–377.
Ward, J.H., Jr. 1963. Hierarchical grouping to optimize an objective function. J. Amer. Stat. Assoc. 58(301):236–244, doi: 10.1080/01621459.1963.10500845.