bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Title: Plant architecture is correlated with variation in cannabinoid concentration and biomass production in *Cannabis sativa*

George M. Stack[1], Craig H. Carlson[1,†], Jacob A. Toth[1], Glenn Philippe[2], Jamie L. Crawford[3], Julie L. Hansen[3], Donald R. Viands[3], Jocelyn K. C. Rose[2], Lawrence B. Smart[1,§]

[1] Horticulture Section, School of Integrative Plant Science, Cornell University, Cornell AgriTech, Geneva, NY
[2] Plant Biology Section, School of Integrative Plant Science, Cornell University, Ithaca, NY
[3] Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, NY
† Current affiliation: Cereal Crops Research Unit, Edward T. Schafer Agricultural Research Center, USDA-ARS, Fargo, ND

[§]Corresponding Author:
Lawrence B. Smart
630 West North St.
Geneva, NY, 14456
Email: lbs33@cornell.edu
Phone: 315-787-2490

Author Contact Information:
GMS: gms252@cornell.edu
CHC: craig.h.carlson@usda.gov
JAT: jat363@cornell.edu
GP: gp376@cornell.edu
JLC: jln15@cornell.edu
JLH: jlh17@cornell.edu
DRV: drv3@cornell.edu
JKCR: jr286@cornell.edu

Submission Date: 12/1/2022
Tables: 1
Figures: 7
Supplementary Tables: 2
Supplementary Figures: 2
Supplementary Datasets: 3
Word Count: 3190

Running Title: Intra-plant cannabinoid variation in *C. sativa*

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Highlight

In-season hemp plant architecture measurements can predict post-harvest traits related to the distribution of biomass and concentration of cannabinoids.

Abstract

Cannabis sativa is cultivated for multiple uses including the production of cannabinoids. In developing improved production systems for high-cannabinoid cultivars, scientists and cultivators must consider the optimization of complex and interacting sets of morphological, phenological, and biochemical traits, which have historically been shaped by natural and anthropogenic selection. Determining factors that modulate cannabinoid variation within and among genotypes is fundamental to developing efficient production systems and understanding the ecological significance of cannabinoids. Thirty-two high-cannabinoid hemp cultivars were characterized for traits including flowering date and shoot-tip cannabinoid concentration. Additionally, a set of plant architecture traits, as well as wet, dry, and stripped inflorescence biomass were measured at harvest. One plant per plot was partitioned post-harvest to quantify intra-plant variation in inflorescence biomass production and cannabinoid concentration. Some cultivars showed intra-plant variation in cannabinoid concentration, while many had a consistent concentration regardless of canopy position. There was both intra- and inter-cultivar variation in architecture that correlated with intra-plant distribution of inflorescence biomass, and concentration of cannabinoids sampled from various positions within a plant. These relationships among morphological and biochemical traits will inform future decisions by cultivators, regulators, and plant breeders as well as our broader understanding of intra-plant variation of specialized biochemicals.

Keywords: Cannabinoids, cannabis, cannabidiol, canopy architecture, hemp, intra-plant variation, secondary metabolites

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Abbreviations

Δ^9-tetrahydrocannabinol (THC)

Tetrahydrocannabinolic acid (THCA)

Cannabidiol (CBD)

Cannabidiolic acid (CBDA)

Cannabichromenic acid (CBCA)

Cannabichromene (CBC)

Cannabigerolic acid (CBGA)

Cannabigerol (CBG)

Cannabinol (CBN)

Tetrahydrocannabivarin (THCV)

Tetrahydrocannabivarinic acid (THCVA)

Cannabidivarin (CBDV)

Cannabidivarinic acid (CBDVA)

Cannabicyclol (CBL)

Cannabicyclolic acid (CBLA)

Δ^8-tetrahydrocannabinol (Δ^8-THC)

Maximum canopy diameter (MCD)

Height at maximum canopy diameter (MCDH)

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

1 Introduction

Plant secondary metabolism generates diverse classes of chemical compounds that have numerous ecological and anthropic functions. Many of these compounds mediate interactions between plants and their environment (Bennett and Wallsgrove, 1994; Wink, 2008; Kessler and Kalske, 2018), and more recent studies have demonstrated their complex multifunctionality and utility in primary metabolic processes (Erb and Kliebenstein, 2020). Ecological theory predicts ontogenetic and tissue-specific variation in the concentration of plant secondary metabolites, primarily in the context of herbivore resistance (McKey, 1974; Pavia *et al.*, 2002; van Dam, 2009; Meldau *et al.*, 2012; Schuman and Baldwin, 2016; Barton and Boege, 2017). There are many examples of ecologically relevant intra-organismal variation in secondary metabolite concentration across diverse taxa, including *Gossypium* (Anderson and Agrell, 2005), *Brassica* (Gutbrodt *et al.*, 2012), *Acer* (Mason *et al.*, 2019), *Populus* (Mason *et al.*, 2019), *Caulerpa* (Meyer and Paul, 1992), *Dictyota* (Cronin and Hay, 1996), and *Oceanapia* (Schupp *et al.*, 1999). The well-studied cases of biochemically-mediated plant resistance to herbivores provide insight into the evolutionary drivers responsible for heritable intra-plant variation in secondary metabolites.

Cannabis sativa synthesizes and stores cannabinoids, a class of secondary metabolites, in glandular trichomes that are most densely produced on female inflorescences (Dayanandan and Kaufman, 1976; Turner *et al.*, 1978; Livingston *et al.*, 2020). The adaptive value of cannabinoid synthesis prior to human cultivation is not known, but theories include protection from herbivores, pathogens, or ultraviolet radiation (Pate, 1983; Lydon *et al.*, 1987; Gorelick and Bernstein, 2017; Tanney *et al.*, 2021). In any of these cases, one might predict that intra-plant variation in cannabinoid production, with greater concentrations of cannabinoids in tissues with a greater fitness value, would be evolutionarily advantageous, as cannabinoids are costly to synthesize. Consistent with this prediction, Richins et al. (2018) and Bernstein et al. (2019) reported significantly greater concentrations of cannabinoids in inflorescences than in vegetative leaves; however, this may be, in part, a result of strong human-driven selection for cannabinoid-rich inflorescences.

Beyond the hypothesized adaptive value, cannabinoid concentration is among the most important factors considered by commercial producers and processors of high-cannabinoid *C. sativa*. The two major harvested products of high-cannabinoid *C. sativa* are bulk inflorescence biomass for

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

extraction and trimmed smokable flowers for direct consumption. For producers of high-cannabinoid hemp, the concentrations of cannabinoids, specifically Δ^9-tetrahydrocannabinol (THC) and tetrahydrocannabinolic acid (THCA), are of particular importance in the context of regulatory compliance (USDA-AMS, 2021). In most jurisdictions, there is a legislated threshold of total potential THC concentration that distinguishes hemp from marijuana, and if exceeded, growers risk losing both their crop and license to grow hemp.

The samples for regulatory testing for THC compliance and 'potency' testing to determine the value of biomass or smokable flower are collected at various time points of the crop production and post-harvest processing pipeline. Regulatory tissue samples in hemp production are typically taken from the top of an inflorescence several weeks prior to harvest, but protocols vary by jurisdiction. Biomass is often sampled after drying and stripping the floral material from stems and branches. However, this process is often imprecise, and the low-cannabinoid content of residual stem and leaf biomass can dilute the concentration in a bulk sample. Smokable flower samples are often tested after the inflorescence has been trimmed of leaf tissue, leaving predominately trichome-rich perigonal bracts.

Previous studies have demonstrated that some *C. sativa* genotypes exhibit intra-plant variation in cannabinoid concentration. Richins et al. (2018) and Bernstein et al. (2019) both quantified cannabinoids from various regions of the canopy, finding generally greater concentrations of cannabinoids in samples from the upper canopy than from the lower canopy. These findings align with anecdotal evidence from high-cannabinoid *C. sativa* producers who consider the primary apical inflorescence as being the most cannabinoid-rich. Additionally, many of these producers impose labor-intensive plant architecture manipulations through pruning with the goal of increasing inflorescence biomass production and cannabinoid concentration, while mitigating biotic pressures.

The interaction between canopy architecture and canopy position on cannabinoid concentration has recently been investigated by measuring the effects of various pruning treatments. Crispim Massuela et al. (2022) sampled inflorescences from three canopy positions (upper, middle, and lower) of plants that were un-pruned, had apical meristems cut back, or had the lower canopy branches removed. They found that only inflorescence position, not pruning technique or its interaction with inflorescence position, had a significant effect on total cannabidiol (CBD) concentration. Danziger and Bernstein (2021) imposed eight different pruning treatments

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

consisting of various combinations of leaf and branch removal from various parts of the canopy, and computed a 'plant uniformity score' to quantify deviation from the mean concentration of cannabinoids among inflorescences. They found that pruning treatments can increase the computed plant uniformity score through small cannabinoid concentration increases in inflorescences that were shaded in the un-pruned control. Pruning either decreased or did not change the concentration of total THC, the predominate cannabinoid, in the primary apical inflorescence. These studies provide important insight into the effect of canopy architecture on intra-plant variation in cannabinoid production. However, additional related studies of a comparatively more diverse set of *C. sativa* germplasm are needed to validate these findings, as well as account for potential genotype-by-architecture interactions.

The primary objectives of this study were to (1) characterize a set of high-CBD hemp cultivars for in-season morphological and biochemical traits, (2) characterize the partitioning of stripped inflorescence biomass throughout the plant canopy, (3) quantify intra-plant variation in cannabinoid concentration, and (4) evaluate the relationships between morphological and biochemical traits to inform ecological models, as well as future decisions by hemp cultivators, regulators, and breeders.

2 Materials and Methods

2.1 Plant material and propagation

Thirty-two high-CBD hemp cultivars (Table 1) from nine commercial sources and the Cornell hemp breeding program were established in a peat-based soilless media (Lambert LM-111) in the second week of May 2020. In accordance with techniques employed by commercial cultivators, plants were propagated from dioecious (male and female) seed, 'feminized' (all female) seed, or via two-node vegetative cuttings. Dioecious cultivars were screened at the seedling stage with the *C. sativa* Y chromosome-specific molecular marker, CSP-1 (Toth *et al.*, 2020), in order to select exclusively females for transplant to the field. Cuttings were rooted using Clonex® rooting hormone (Hydrodynamics International, Lansing, MI). Seedlings and cuttings were maintained in the greenhouse at 18 h light:6 h dark until transplant in the first week of June.

2.2 Field preparation and maintenance

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Trials were planted at two Cornell University field sites: Geneva, NY (McCarthy Farm: 42.9, -77.0) and Ithaca, NY (Bluegrass Lane Turf and Ornamental Farm: 42.5, -76.5). Weather station data can be found at http://newa.cornell.edu (Last accessed on 11/22/2022). Each site was cultivated and raised beds formed with drip irrigation and black plastic mulch prepared every 1.83 m on center. Fertilizer (19-19-19, Phelps Supply Inc., Phelps, NY) equivalent to 95.3 kg N ha^{-1}, was spread under the plastic mulch in Geneva and was broadcast pre-planting in Ithaca. Landscape fabric was deployed to control weeds in alleys between rows.

Each cultivar was planted in five-plant plots in a replicated, randomized complete block design with four replicate blocks at each site. Seedlings and rooted cuttings were transplanted into raised beds on June 11, 2020 (Geneva, NY) and June 12, 2020 (Ithaca, NY). Plants were spaced 1.23 m apart within rows. After transplanting, the plots were irrigated using in-bed drip irrigation as needed throughout the season to maintain optimal soil moisture (>0.27 m^3 m^{-3}). HOBOnet RXW-SMD-10HS soil moisture sensors (Onset, Bourne, MA) were installed at a depth of 10 cm and used to assess when irrigation was necessary. Fertilizer (Jack's 12-4-16 Hydro FeED RO, 11.3 kg per application) was included in the irrigation on two occasions in early and late July.

2.3 In-season measurements

The heights of the middle three plants of each five-plant plot were measured weekly after transplant until there was no change in height for two consecutive weeks. All plants were surveyed weekly to determine the onset of flowering following the protocol described by Carlson et al. (2021), where 'terminal flowering' describes distinct clusters of pistillate flowers observed at shoot apices (Supplementary Fig. S1). Plants that produced staminate flowers were promptly removed from the field to maintain unpollinated female inflorescences.

For cannabinoid quantification, regulatory-style shoot-tip samples of the 32 cultivars were collected from a single plant in every plot starting one week after terminal flowering and again at three and five weeks after terminal flowering. In accordance with regulatory standard in New York State at the time of the trial, the top 10 cm of the apical inflorescence was sampled for the time series. The first plant in the plot was sampled for the 1-week timepoint, the fifth for the 3-week timepoint, and the second for the 5-week timepoint. Shoot tip samples were dried in a climate-controlled room (~30% RH and below 33°C) and then milled to a fine powder in a Ninja® Pro blender (SharkNinja, Needham, MA). Milled samples were stored at 4°C prior to

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

high pressure liquid chromatography (HPLC) analysis following the methods described by Stack et al. (2021) (Supplementary Dataset S2). The following cannabinoids were quantified for each sample: tetrahydrocannabinolic acid (THCA), Δ^9-tetrahydrocannabinol (THC), cannabidiolic acid (CBDA), cannabidiol (CBD), cannabichromenic acid (CBCA), cannabichromene (CBC), cannabigerolic acid (CBGA), cannabigerol (CBG), cannabinol (CBN), tetrahydrocannabivarin (THCV), tetrahydrocannabivarinic acid (THCVA), cannabidivarin (CBDV), cannabidivarinic acid (CBDVA), cannabicyclol (CBL), cannabicyclolic acid (CBLA), and Δ^8-tetrahydrocannabinol (Δ^8-THC). To control for variation in decarboxylation of acid-form cannabinoids and cultivar-dependent variation in relative proportions of different cannabinoids, the statistical analysis was performed on the sum of all total potential cannabinoid percentages. See Table 3 in Stack et al. (2021) for formulas used to calculate total potential cannabinoids.

2.4 Time-of-harvest and post-harvest measurements

At harvest, 5 weeks after terminal flowering, canopy architecture traits of the third plant in each plot were measured following the method described by Carlson et al. (2021), including height, maximum canopy diameter (MCD), and the height at the maximum canopy diameter (MCDH). Following the measurements, the stems were cut at soil level and the total wet biomass of each plant in a plot was measured. The third plant in each five-plant plot was partitioned into five sections (S1 through S5) based on the length of the main stem (Fig. 1). Sections were subsequently air-dried in a greenhouse with industrial fans.

After drying, basal stem diameter, total dry biomass – including stems, leaves, and inflorescence material – and dry stripped inflorescence biomass were measured for each section. The full list of traits measured with relevant formulas can be found in Supplementary Table S1. Cannabinoids were quantified by HPLC following the protocol above using a subsample of the homogenized stripped biomass for each section (Supplementary Dataset S3).

2.7 Statistical analysis

All statistical analyses and preparation of graphs were conducted in the open-source statistical computing platform R version 4.1.3 (R Core Team, 2022). Simple linear regression was used to determine cultivar-wise correlation and regression coefficients among cannabinoid sampling timepoints and positions. The cultivar-level mean proportions of dry biomass in each section

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

were used to conduct a *k*-means clustering analysis following the Hartigan and Wong algorithm (Hartigan and Wong, 1979). The number of clusters (n=4) was determined by the elbow method (Supplementary Fig. S2). Simple linear regression was used at both the cultivar- and plot-level to determine correlations between the following values: the proportion of dry biomass in S5 and stripped inflorescence biomass per unit area, the MCD to height ratio and the proportion of dry biomass in S5, the MCD to height ratio and the log of the ratio between stripped inflorescence biomass produced in S4 and S5. Second-degree polynomial regression was used to model cultivar- and plot-level relationships between bicone volume and harvest index.

A mixed-effects model with biomass section and cultivar as fixed effects, and plot as a random effect, was used to determine whether a cultivar or biomass section had a significant effect on total cannabinoid concentration. A series of mixed-effects model F-tests were used to determine whether biomass section had a significant effect on total cannabinoid concentration for each cultivar. For all mixed-effects models a Satterthwaite approximation was used to estimate the effective degrees of freedom. When the effect of section was deemed significant after correcting for multiple testing using a Bonferroni correction ($\alpha = 0.0016$), a post-hoc Tukey's HSD analysis was used to test pairwise differences between sections.

To follow-up on the results of previous studies that reported intra-plant variation in cannabinoid concentration, cultivar-level cannabinoid concentration was regressed on plant section to calculate a slope corresponding to the change in cannabinoid concentration between sections from the top and moving down each section. Based on the results of Danziger and Bernstein (2021), canopy area, dry canopy density, and an interaction term were used to model the change in cannabinoid content by section at the cultivar level. The dry canopy density was log transformed to normalize the residuals, and non-significant terms were removed from the model before reporting model statistics.

Stepwise regression was performed with the function 'step' (direction='both') using all biomass and plant architecture variables found in Supplementary Dataset S1 to model change in cannabinoid concentration by section at the cultivar-level. The package relaimpo (Groemping, 2007) was used to order predictors and calculate LMG indices to partition additive properties of R^2. A mixed-linear model was used to predict change in cannabinoid concentration by section using the log of dry canopy density, canopy area, and an interaction term as predictors.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

3 Results

3.1 Cannabinoid concentrations in regulatory samples are well correlated with concentrations in end-of-season biomass

Total potential cannabinoid concentration was positively correlated among all in-field sampling timepoints and postharvest biomass samples (Fig. 2, $p < 0.05$). The correlations between regulatory-style shoot samples and stripped biomass samples were strong, with R^2 values ranging from 0.64 to 0.85. The concentration in end-of-season stripped bulk biomass was more strongly correlated with the concentration in samples from the biomass sections than the concentration in regulatory-style shoot tip samples. Samples from biomass sections generally had a greater correlation coefficient with those from adjacent sections, rather than with those from more distal sections. Samples from the bulk biomass had a greater correlation coefficient with lower sections of the canopy than with samples from upper sections.

3.2 Biomass distribution among sections is cultivar-dependent

The cultivars evaluated produced a wide range of stripped inflorescence biomass on a per-plant basis (Fig. 3A). Further, the proportion of total stripped biomass produced in each section of the canopy varied by cultivar (Fig. 3B). On one extreme, over 75% of the total stripped biomass produced by GVA-H-19-1077-008 was in S5, the section closest to the base of the plant. In contrast, less than 30% of the total biomass produced by 'Umpqua' was in S5. *K*-means clustering analysis of the proportions of dry stripped biomass in each section separated the cultivars into four clusters (Fig. 3B). The clusters corresponded almost perfectly with the variation in the proportion of stripped biomass in S5.

3.3 Variation in plant architecture is correlated with biomass distribution

There was substantial variation in plant architecture among and within the 32 cultivars evaluated with plant height and MCD ranging from 61 cm to 249 cm and 55 cm to 265 cm, respectively (Fig. 4, Supplementary Dataset S1). The ratio of the MCD to height was positively correlated with the proportion of dry biomass in S5 at the level of cultivar ($R^2 = 0.47$, $F(1,30) = 26.1$, $p < 0.001$) and plot ($R^2 = 0.29$, $F(1,228) = 91.35$, $p < 0.001$) (Fig. 5), and negatively correlated with the log of the ratio between stripped biomass in S4 and S5 at the level of cultivar ($R^2 = 0.44$, $F(1,30) = 23.37$, $p < 0.001$) and plot ($R^2 = 0.20$, $F(1,228) = 55.28$, $p < 0.001$) (Fig. 5). There

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

was a significant relationship between bicone volume and harvest index at the level of cultivar ($R^2 = 0.57$, $F(2,29) = 55.28$, $p < 0.001$) and plot ($R^2 = 0.39$, $F(2,227) = 73.11$, $p < 0.001$) (Fig. 4). The proportion of dry biomass in S5 was negatively correlated with the log of the amount of stripped inflorescence biomass per unit area at the level of cultivar ($R^2 = 0.33$, $F(1,30) = 14.56$, $p < 0.001$) and plot ($R^2 = 0.09$, $F(1,228) = 23.81$, $p < 0.001$) (Fig 5).

3.4 Intra-plant variation in cannabinoid concentration is cultivar-dependent

When modeling the concentration of total potential cannabinoids sampled from the biomass sections using a mixed-effects model, there were significant main effects of cultivar ($F(31,502.12) = 16.23$, $p < 0.001$) and biomass section ($F(1,988.23) = 434.14$, $p < 0.001$), and a significant interaction between cultivar and biomass section ($F(31,988.22) = 5.24$, $p < 0.001$). After correcting for multiple testing for the 32 mixed-effects model F-tests ($\alpha = 0.0016$), 17 of the cultivars had statistically significant variation in total cannabinoid concentration among plant sections (Fig. 6). For those cultivars with significant variation among sections, the cannabinoid concentration in S1 was always significantly greater than the concentration in S5.

3.5 Variations in plant biomass and architecture are correlated with intra-plant variation in cannabinoid concentration

In the multiple linear regression predicting change in cannabinoid content by plant section, the overall model, after the main effect of dry canopy density was removed, was significant ($R^2 = 0.30$, $F(2,29) = 6,26$, $p = 0.006$) (Fig. 7A). Both canopy area ($\beta = 1.460$, $p = 0.006$) and the interaction between canopy area and the log of dry canopy density ($\beta = -0.273$, $p = 0.003$) were found to be significant predictors of the change in cannabinoid content by plant section.

To predict the change in cannabinoid content by plant section using stepwise regression, nine predictors were identified that explained 27.7% of the variance (Fig. 7B). The regression statistics and LMG values for the nine predictors can be found in Supplementary Table S2. Based on the LMG metric, the amount of stripped inflorescence biomass in S4 was the best single predictor, explaining 8.1% of the variance.

4 Discussion

4.1 Correlations between in-season and post-harvest cannabinoid concentrations

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

The strong correlation between cannabinoid concentration early in inflorescence development and concentrations in bulk stripped biomass suggests that the metabolic mechanism modulating variation in cannabinoid concentration among cultivars is initiated soon after the onset of flowering. This trend may persist earlier into floral development, with tissue accumulating greater concentrations of cannabinoids after floral induction, but before the visible onset of flowering. Richins et al. (2018) reported a significant correlation between THC concentrations in vegetative fan leaves and inflorescence material ($R^2 = 0.358$) even before the induction of flowering.

While there is a substantial and growing body of data concerning cannabinoid accumulation after the visible onset of flowering (Pacifico *et al.*, 2008; De Backer *et al.*, 2012; Aizpurua-Olaizola *et al.*, 2016; Richins *et al.*, 2018; Yang *et al.*, 2020; Stack *et al.*, 2021), far less is known about variation and accumulation of cannabinoids prior to this developmental transition. With a better understanding of intra-plant cannabinoid variation prior to flowering, and the relationships between pre-inflorescence and at-harvest cannabinoid concentrations, we could broaden our mechanistic understanding of cannabinoid biosynthesis and could uncover powerful early-cycle selection criteria to breed improved cultivars. Additionally, the strong correlations between samples throughout the course of floral development could be used as a predictive tool for high-cannabinoid hemp producers to schedule regulatory testing while remaining compliant for THC concentration.

4.2 Distribution of inflorescence biomass within the plant canopy and relationships with plant architecture

As inflorescence biomass is the primary product of interest when producing high-cannabinoid *C. sativa*, the position and distribution of this biomass throughout the canopy is critical for growers. Based on the data presented here, growers should prefer plants that produce a greater proportion of biomass in the upper canopy. Plants that had a greater proportion of biomass in S5 produced less inflorescence biomass per unit area (Fig. 5C), and plants with a large bicone volume also had a significantly lower harvest index, potentially due to the additional stem biomass needed to support the large basal branches (Fig. 4B).

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Moving toward a high-density production system, where individual plants are smaller but production per unit area is greater, has been effective in increasing yields across dozens of cropping systems (Duvick, 2005; Assefa *et al.*, 2018; Postma *et al.*, 2021). Through tandem improvement of cultural practices for high-density cultivation, as well as development of cultivars selected for high-density production, cultivators have been able to increase yields and production efficiency. While the bulk of this discussion focuses on outdoor production systems, similar principles govern increasing production density in controlled environment systems (Rodriguez *et al.*, 2007; Amundson *et al.*, 2012). Selection strategies towards improved high-cannabinoid cultivars for indoor production should consider optimization of biomass distribution for high-density cultivation systems, specifically by reducing the need for labor-intensive manual pruning and training.

4.3 Cultivar-dependent intra-plant variation in cannabinoid concentration

In contrast with previous studies of intra-plant variation in cannabinoid concentration (Richins *et al.*, 2018; Bernstein *et al.*, 2019; Danziger and Bernstein, 2021; Crispim Massuela *et al.*, 2022), many of the cultivars evaluated in this study had little variance in cannabinoid concentration across the canopy. This incongruence could be attributed to many factors, including the different genotypes used in the studies, variation in sampling methodology, and differences in cultivation environment. Nevertheless, all cultivars with significant variation among sections followed the same general pattern described in the literature, with the greatest concentrations of cannabinoids accumulating in the upper-most part of the canopy.

The degree of variation in total cannabinoid concentration throughout the canopy is particularly important to growers in the context of regulatory compliance. Because THCA and CBDA are synthesized in predictable ratios by the CBDA synthase enzyme in chemotype III *C. sativa* (Zirpel *et al.*, 2018; Toth *et al.*, 2020, 2021; Yang *et al.*, 2020; Stack *et al.*, 2021), remediation of non-compliant floral biomass by blending it with plant tissue containing lower THC concentration is effective primarily due to differences in the total concentration of cannabinoids among plant sections. If there is not substantial variation in cannabinoid concentration in stripped inflorescence biomass by section, dilution with inflorescence biomass from lower sections of the canopy would not reduce the total THC concentration sufficiently to bring it below the compliance threshold.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

4.4 Correlations with intra-plant variation in cannabinoid concentration

The factors underpinning the variation, or lack of variation, in cannabinoid concentration throughout the canopy must be understood to develop efficient production systems. This dataset complements previous work investigating the impact of plant architecture manipulation on intra-plant variation in cannabinoid accumulation (Danziger and Bernstein, 2021; Crispim Massuela *et al.*, 2022). Specifically, given that manipulation of plant architecture and canopy density can cause local changes in cannabinoid concentration, does existing variation in plant architecture among cultivars correlate with intra-plant variation in cannabinoid content? Interestingly, there was a significant interaction between canopy area and dry canopy density in predicting decrease in cannabinoid concentration. Plants with a small canopy area, independent of canopy density, showed very little reduction in cannabinoid concentration, while plants with a greater canopy area exhibited canopy density-dependent decreases in cannabinoid concentration. This lends support to hypotheses that factors including light penetration and density of non-inflorescence leaves influence intra-plant variation in cannabinoid concentration. In many species, variation in light penetration has been shown to contribute to intra-plant variation in secondary metabolites, including: flavonoids (Del Valle *et al.*, 2018), anthocyanins (González-Talice *et al.*, 2013), and tannins (Mole *et al.*, 1988). Additional work examining light penetration into the canopy, photosynthetic activity throughout the canopy, and sub-canopy lighting will provide a better understanding of the degree to which variability in light intensity modulates variation in cannabinoid concentration.

Stepwise regression modeling indicated that stripped inflorescence biomass produced in S4 was the best single predictor of intra-plant variation cannabinoid content, such that plants with greater amounts of stripped inflorescence biomass in S4 had a greater decrease in cannabinoid concentration moving down the plant. The next most important predictors were the amount of biomass in S3, the proportion of S4 that comprised stripped inflorescence biomass, and the kite angle ratio. Of the nine predictors in the final model, six were biomass measurements, while the other three were plant architecture measurements. Though it was not selected as a predictor in the final model, the relationships between the selected plant architecture and biomass measurements with canopy density suggest that variation in micro-environments throughout the plant canopy drives variation in cannabinoid concentration. Plants that produce less stripped

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

inflorescence biomass and have a lower relative MCDH could have a more homogenous and open canopy environment (light, temperature, humidity, airflow, etc.) and thus a more homogeneous cannabinoid profile throughout the canopy.

Further studies are needed to characterize the interactions between genotype, environment, planting density, and manipulation of plant architecture, specifically contrasting genotypes with and without innate intra-plant variation in cannabinoid concentration.

4.5 Is intra-canopy variation in cannabinoid concentration adaptive?

Based on this study and others, intra-plant variation in cannabinoid concentration correlates with variability in photosynthate production driven by microclimates within the canopy. However, there are alternative, and not mutually exclusive, hypotheses that could explain intra-plant variation in cannabinoids. For example, assuming that cannabinoids confer protection to proximal tissues, adaptive variation in cannabinoid concentration among inflorescences from various positions within the canopy is consistent with optimal defense theory (Pavia *et al.*, 2002). Cannabinoid synthesis is concentrated in tissues proximal to those producing seeds, and as such may be under strong selective pressure through influencing seed survival, dispersal, and subsequent reproductive success contributing to plant fitness. It is well documented in many angiosperms that seeds produced in different regions of the canopy can vary in size and biochemical composition (Hendrix, 1984; Raboy and Dickinson, 1987; Huber *et al.*, 2016). Further, intra-specific variation in seed size is widely associated with differences in seed dispersal and seeding establishment (Stanton, 1984; Baskin and Baskin, 2014; Visser *et al.*, 2016; Genna and Pérez, 2021). Thus, variation in seed provisioning and fitness throughout the *C. sativa* canopy, potentially driven by microclimatic variation, could drive natural selection for variation in cannabinoid production among inflorescences, with greater concentrations of cannabinoids being produced in inflorescences with better-provisioned seeds. Seeds and cannabinoids are both metabolic sinks and their proximity and strength relative to photosynthetic source tissues likely limits their production. This study could not capture variation in seed traits as the plants were maintained unpollinated. Little is known about intra-plant variation in seed traits in *C. sativa* and future studies, particularly studying wild or feral germplasm, which has been under natural selection, would be valuable in the development of ecological models.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Broader application of these models to other species might help to dissect the basis of adaptive intra-plant variation in secondary metabolites.

5 Supplementary Data

Table S1. List of traits and formulas used in the study.

Table S2. Analysis of variance from stepwise selection as well as relative importance (R %) of predictor variables for change in cannabinoid concentration by section (n=226).

Fig. S1. Flowering stage rating scale for pistillate *C. sativa* following the descriptors of Carlson et al. (2021).

Fig. S2. Elbow plot for *k*-means clustering of cultivars by dry biomass proportions.

Dataset S1. Plot-level summary statistics for all traits.

Dataset S2. HPLC data showing cannabinoid concentrations from 1w, 3w, and 5w regulatory shoot tip samples.

Dataset S3. HPLC data showing cannabinoid concentrations from biomass section subsamples.

6 Acknowledgements

We are grateful to the research teams from the L. Smart, C. Smart, Viands, Rose, and Bergstrom labs for their excellent technical assistance on this project, especially Deanna Gentner, Stephen Snyder, Allison DeSario, Teagan Zingg, Savanna Shelnutt, Emily McFadden, Lauren Carlson, Colin Day, Garrett Giles, Dr. Heather Grab, Nicholas Kaczmar, Jesse Chavez, Ryan Crawford, Jason Schiller, Johanna Gertin and the farm crews at Cornell AgriTech and the CUAES campus area farms. We are also grateful to Stem Holdings Agri, Sunrise Genetics, Phytonyx, Industrial Seed Innovations, Castetter Sustainability Group, Ryes Creek, Front Range Biosciences, Charlotte's Web, Blue Forest Farms, Wessels' Farm, and Boring Hemp for generously contributing cultivars.

7 Author Contributions

GMS: conceptualization, data curation, formal analysis, investigation, methodology, writing – original draft preparation. CHC, JAT: investigation, methodology. JLC: investigation. GP: methodology. JKCR, DRV, JLH, LBS: funding acquisition, supervision. All authors contributed to writing – review and editing.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

8 Conflicts of Interest

The authors declare no conflicts of interest.

9 Funding

This trial was funded by a sponsored research agreement with Pyxus International and by grants (AC477 and AC483) from New York State Dept. of Agriculture and Markets through the Empire State Development Corporation. This work was supported in part by the U.S. Department of Agriculture, Agricultural Research Service through project 3060-21000-038-000-D.

10 Data Availability

All data supporting the findings of this study are available within the paper and within its supplementary materials published online.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

11 References

Aizpurua-Olaizola O, Soydaner U, Öztürk E, Schibano D, Simsir Y, Navarro P, Etxebarria N, Usobiaga A. 2016. Evolution of the cannabinoid and terpene content during the growth of *Cannabis sativa* plants from different chemotypes. Journal of Natural Products **79**, 324–331.

Amundson S, Deyton DE, Kopsell DA, Hitch W, Moore A, Sams CE. 2012. Optimizing plant density and production systems to maximize yield of greenhouse-grown 'Trust' tomatoes. HortTechnology **22**, 44–48.

Anderson P, Agrell J. 2005. Within-plant variation in induced defence in developing leaves of cotton plants. Oecologia **144**, 427–434.

Assefa Y, Carter P, Hinds M, Bhalla G, Schon R, Jeschke M, Paszkiewicz S, Smith S, Ciampitti IA. 2018. Analysis of long term study indicates both agronomic optimal plant density and increase maize yield per plant contributed to yield gain. Scientific Reports **8**, 4937.

Barton KE, Boege K. 2017. Future directions in the ontogeny of plant defence: understanding the evolutionary causes and consequences. Ecology Letters **20**, 403–411.

Baskin CC, Baskin JM. 2014. *Seeds: ecology, biogeography and evolution of dormancy and germination*. San Diego, CA, USA: Academic Press.

Bennett RN, Wallsgrove RM. 1994. Secondary metabolites in plant defence mechanisms. The New Phytologist **127**, 617–633.

Bernstein N, Gorelick J, Koch S. 2019. Interplay between chemistry and morphology in medical cannabis (*Cannabis sativa* L.). Industrial Crops and Products **129**, 185–194.

Carlson CH, Stack GM, Jiang Y, Taşkıran B, Cala AR, Toth JA, Philippe G, Rose JKC, Smart CD, Smart LB. 2021. Morphometric relationships and their contribution to biomass and cannabinoid yield in hybrids of hemp (*Cannabis sativa*). Journal of Experimental Botany **72**, 7694–7709.

Crispim Massuela D, Hartung J, Munz S, Erpenbach F, Graeff-Hönninger S. 2022. Impact of harvest time and pruning technique on total CBD concentration and yield of medicinal cannabis. Plants **11**.

Cronin G, Hay ME. 1996. Within-plant variation in seaweed palatability and chemical defenses: optimal defense theory versus the growth-differentiation balance hypothesis. Oecologia **105**, 361–368.

van Dam NM. 2009. Belowground herbivory and plant defenses. Annual Review of Ecology, Evolution, and Systematics **40**, 373–391.

Danziger N, Bernstein N. 2021. Shape matters: plant architecture affects chemical uniformity in large-size medical cannabis plants. Plants **10**, 1834.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Dayanandan P, Kaufman PB. 1976. Trichomes of *Cannabis sativa* L. (Cannabaceae). American Journal of Botany **63**, 578–591.

De Backer B, Maebe K, Verstraete AG, Charlier C. 2012. Evolution of the content of THC and other major cannabinoids in drug-type cannabis cuttings and seedlings during growth of plants: Evolution of major cannabinoids content during growth of plants. Journal of Forensic Sciences **57**, 918–922.

Del Valle JC, Buide ML, Whittall JB, Narbona E. 2018. Phenotypic plasticity in light-induced flavonoids varies among tissues in *Silene littorea* (Caryophyllaceae). Environmental and Experimental Botany **153**, 100–107.

Duvick DN. 2005. The contribution of breeding to yield advances in maize (*Zea mays* L.). Advances in Agronomy. Academic Press, 83–145.

Erb M, Kliebenstein DJ. 2020. Plant secondary metabolites as defenses, regulators, and primary metabolites: the blurred functional trichotomy. Plant Physiology **184**, 39–52.

Genna NG, Pérez HE. 2021. Intra-population seed survival during burial on a mass basis. Flora **274**, 151756.

González-Talice J, Yuri JA, del Pozo A. 2013. Relations among pigments, color and phenolic concentrations in the peel of two Gala apple strains according to canopy position and light environment. Scientia Horticulturae **151**, 83–89.

Gorelick J, Bernstein N. 2017. Chemical and physical elicitation for enhanced cannabinoid production in *Cannabis*. *Cannabis sativa* L. - Botany and Biotechnology. New York, NY, USA: Springer International Publishing, 439–456.

Groemping U. 2007. Relative importance for linear regression in R: The package relaimpo. Journal of Statistical Software **17**, 1–27.

Gutbrodt B, Dorn S, Unsicker SB, Mody K. 2012. Species-specific responses of herbivores to within-plant and environmentally mediated between-plant variability in plant chemistry. Chemoecology **22**, 101–111.

Hartigan JA, Wong MA. 1979. Algorithm AS 136: a K-means clustering algorithm. Journal of the Royal Statistical Society. Series C, Applied Statistics **28**, 100–108.

Hendrix SD. 1984. Variation in seed weight and its effects on germination in *Pastinaca sativa* L. (Umbelliferae). American Journal of Botany **71**, 795–802.

Huber SC, Li K, Nelson R, Ulanov A, DeMuro CM, Baxter I. 2016. Canopy position has a profound effect on soybean seed composition. PeerJ **4**, e2452.

Kessler A, Kalske A. 2018. Plant secondary metabolite diversity and species interactions. Annual Review of Ecology, Evolution, and Systematics **49**, 115–138.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Livingston SJ, Quilichini TD, Booth JK, Wong DCJ, Rensing KH, Laflamme-Yonkman J, Castellarin SD, Bohlmann J, Page JE, Samuels AL. 2020. *Cannabis* glandular trichomes alter morphology and metabolite content during flower maturation. The Plant Journal **101**, 37–56.

Lydon J, Teramura AH, Coffman CB. 1987. UV-B radiation effects on photosynthesis, growth and cannabinoid production of two *Cannabis sativa* chemotypes. Photochemistry and Photobiology **46**, 201–206.

Mason CJ, Long DC, Lindroth RL, Hoover K. 2019. Divergent host plant utilization by adults and offspring is related to intra-plant variation in chemical defences. The Journal of Animal Ecology **88**, 1789–1798.

McKey D. 1974. Adaptive patterns in alkaloid physiology. The American Naturalist **108**, 305–320.

Meldau S, Erb M, Baldwin IT. 2012. Defence on demand: mechanisms behind optimal defence patterns. Annals of Botany **110**, 1503–1514.

Meyer, Paul. 1992. Intraplant variation in secondary metabolite concentration in three species of *Caulerpa* (Chlorophyta: Caulerpales) and its effects on herbivorous fishes. Marine Ecology Progress Series **82**, 249-257.

Mole S, Ross JA, Waterman PG. 1988. Light-induced variation in phenolic levels in foliage of rain-forest plants : I. Chemical changes. Journal of Chemical Ecology **14**, 1–21.

Pacifico D, Miselli F, Carboni A, Moschella A, Mandolino G. 2008. Time course of cannabinoid accumulation and chemotype development during the growth of *Cannabis sativa* L. Euphytica **160**, 231–240.

Pate DW. 1983. Possible role of ultraviolet radiation in evolution of *Cannabis* chemotypes. Economic Botany **37**, 396.

Pavia H, Toth GB, Åberg P. 2002. Optimal defense theory: Elasticity analysis as a tool to predict intraplant variation in defenses. Ecology **83**, 891–897.

Postma JA, Hecht VL, Hikosaka K, Nord EA, Pons TL, Poorter H. 2021. Dividing the pie: A quantitative review on plant density responses. Plant, Cell & Environment **44**, 1072–1094.

Raboy V, Dickinson DB. 1987. The timing and rate of phytic acid accumulation in developing soybean seeds. Plant Physiology **85**, 841–844.

Richins RD, Rodriguez-Uribe L, Lowe K, Ferral R, O'Connell MA. 2018. Accumulation of bioactive metabolites in cultivated medical *Cannabis*. PloS One **13**, e0201119.

Rodriguez JC, Shaw NL, Cantliffe DJ. 2007. Influence of plant density on yield and fruit quality of greenhouse-grown Galia muskmelons. HortTechnology **17**, 580–585.

Schuman MC, Baldwin IT. 2016. The layers of plant responses to insect herbivores. Annual Review of Entomology **61**, 373–394.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Schupp P, Eder C, Paul V, Proksch P. 1999. Distribution of secondary metabolites in the sponge *Oceanapia* sp. and its ecological implications. Marine Biology **135**, 573–580.

Stack GM, Toth JA, Carlson CH, *et al.* 2021. Season□long characterization of high□cannabinoid hemp (*Cannabis sativa* L.) reveals variation in cannabinoid accumulation, flowering time, and disease resistance. Global Change Biology Bioenergy **13**, 546-561.

Stanton ML. 1984. Seed variation in wild radish: Effect of seed size on components of seedling and adult fitness. Ecology **65**, 1105–1112.

Tanney CAS, Backer R, Geitmann A, Smith DL. 2021. Cannabis glandular trichomes: a cellular metabolite factory. Frontiers in Plant Science **12**, 721986.

Toth JA, Smart LB, Smart CD, Stack GM, Carlson CH, Philippe G, Rose JKC. 2021. Limited effect of environmental stress on cannabinoid profiles in high□cannabidiol hemp (*Cannabis sativa* L.). Global Change Biology Bioenergy **13**, 1666–1674.

Toth JA, Stack GM, Cala AR, *et al.* 2020. Development and validation of genetic markers for sex and cannabinoid chemotype in *Cannabis sativa* L. Global Change Biology Bioenergy **12**, 213–222.

Turner JC, Hemphill JK, Mahlberg PG. 1978. Quantitative determination of cannabinoids in individual glandular trichomes of *Cannabis sativa* L. (Cannabaceae). American Journal of Botany **65**, 1103–1106.

USDA-AMS. 2021. Establishment of a domestic hemp production program. Federal Register **86**, 5596–5691.

Visser MD, Bruijning M, Wright SJ, Muller-Landau HC, Jongejans E, Comita LS, Kroon H. 2016. Functional traits as predictors of vital rates across the life cycle of tropical trees. Functional Ecology **30**, 168–180.

Wink M. 2008. Plant secondary metabolism: diversity, function and its evolution. Natural Product Communications **3**, 1205–1216.

Yang R, Berthold EC, McCurdy CR, da Silva Benevenute S, Brym ZT, Freeman JH. 2020. Development of cannabinoids in flowers of industrial hemp (*Cannabis sativa* L.): a pilot study. Journal of agricultural and food chemistry **68**, 6058–6064.

Zirpel B, Kayser O, Stehle F. 2018. Elucidation of structure-function relationship of THCA and CBDA synthase from *Cannabis sativa* L. Journal of Biotechnology **284**, 17–26.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

12 Tables

Table 1. Sources of 32 high-CBD cultivars included in field trials in Geneva and Ithaca. Cultivars were started from seed (dioecious or feminized) or vegetative cuttings. Cultivars were generously contributed by the companies indicated.

Cultivar/ID	Propagation (cutting/seed)	Source
GVA-H-19-1191	Seed (feminized)	Cornell Hemp Program
GVA-H-19-1091	Seed (dioecious)	Cornell Hemp Program
GVA-H-20-1030	Seed (dioecious)	Cornell Hemp Program
GVA-H-19-1064-003	Cutting	Cornell Hemp Program
GVA-H-19-1068-003	Cutting	Cornell Hemp Program
GVA-H-19-1066-001	Cutting	Cornell Hemp Program
GVA-H-19-1067-001	Cutting	Cornell Hemp Program
GVA-H-19-1077-008	Cutting	Cornell Hemp Program
TJ's CBD	Cutting	Stem Holdings Agri
FL 49	Cutting	Sunrise Genetics
FL 58	Cutting	Sunrise Genetics
FL 70	Cutting	Sunrise Genetics
CJ 2	Cutting	Sunrise Genetics
SB 1	Cutting	Sunrise Genetics
Z 25	Cutting	Sunrise Genetics
NS52	Seed (feminized)	Phytonyx
SR-1	Seed (feminized)	Industrial Seed Innovations
Umpqua	Seed (feminized)	Industrial Seed Innovations
Rogue	Seed (feminized)	Industrial Seed Innovations
The Grand	Seed (feminized)	Boring Hemp
CSG Berry Blossom	Seed (feminized)	Castetter Sustainability Group
Sweetened	Seed (feminized)	Ryes Creek
Carolina Dream	Seed (dioecious)	Ryes Creek
BaOx	Seed (dioecious)	Ryes Creek
Hybrid #5	Cutting	Front Range Biosciences
Hybrid #9	Cutting	Front Range Biosciences
Early Pearly	Cutting	Front Range Biosciences
Lindorea	Seed (feminized)	Charlotte's Web
CW EM-18	Seed (feminized)	Charlotte's Web
CW EM-28	Seed (feminized)	Charlotte's Web
CW EM-31	Seed (feminized)	Charlotte's Web
CW EM-73	Seed (feminized)	Charlotte's Web

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

13 Figure Legends

Figure 1. Schematic of (A) a hemp plant in the field, (B) the branching structure of the hemp plant, (C) the canopy-level plant architecture measurements taken prior to harvest, and (D) the sectioning scheme used to partition the plant into five sections.

Figure 2. Cultivar-level pairwise correlations among total potential cannabinoid concentrations from samples over the course of floral maturation and harvest. The main diagonal is the distribution of data for the nine sampling points, from left to right: regulatory-style sample 1 week, 3 weeks, and 5 weeks after terminal flowering, post-harvest stripped biomass samples from plants partitioned into five sections where S1 is branches derived from the top 5th of the main stem, S2 those from the second to top, etc., and bulk inflorescence biomass from the entire plant. Correlation and regression coefficients correspond to the regressions plotted in panels mirrored by the main diagonal.

Figure 3. Bar plots of stripped biomass produced by 32 high-CBD hemp cultivars partitioned into five sections where S1 is branches derived from the top 5th of the main stem, S2 those from the second to top, etc. Colors indicate the mean stripped biomass per plant (A) and the proportion of stripped biomass (B) from each of the five sections. Clusters in (B) are cultivar-level *k*-means clusters based on the proportions of dry biomass produced in each section.

Figure 4. Variation in plant architecture within and among high-cannabinoid hemp cultivars. (A) plot of maximum canopy diameter (MCD) versus height where points are cultivar means and error bars are the standard error of the mean. Dashed lines indicate MCD to height ratios of 0.5, 1, and 2. (B) Second-degree polynomial regression of harvest index on bicone volume where points are cultivar means and error bars are the standard error of the mean. (C) Average kite models by cultivar contrasted from MCD, height, and the height of the maximum canopy diameter (MCDH). Colors in all panels indicate the *k*-means cluster to which that a cultivar belongs based on the proportions of dry biomass in each biomass section.

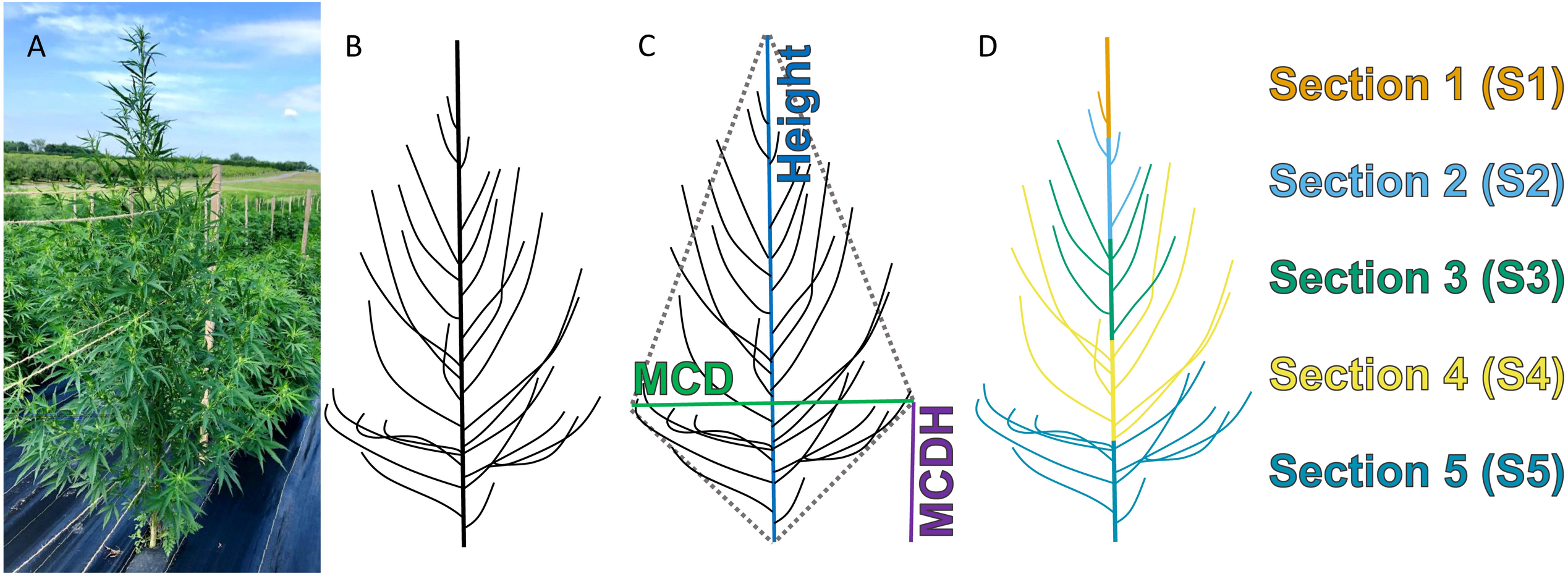
Figure 5. Correlations between (A) MCD:Height ratio and proportion of dry biomass in S5, (B) MCD:Height ratio and log of the ratio of stripped biomass produced in S4 and S5, and (C) proportion of dry biomass in S5 and log of stripped biomass per unit area. Each point and set of error bars, depicting the standard error of the mean, represents one cultivar. Colors indicate

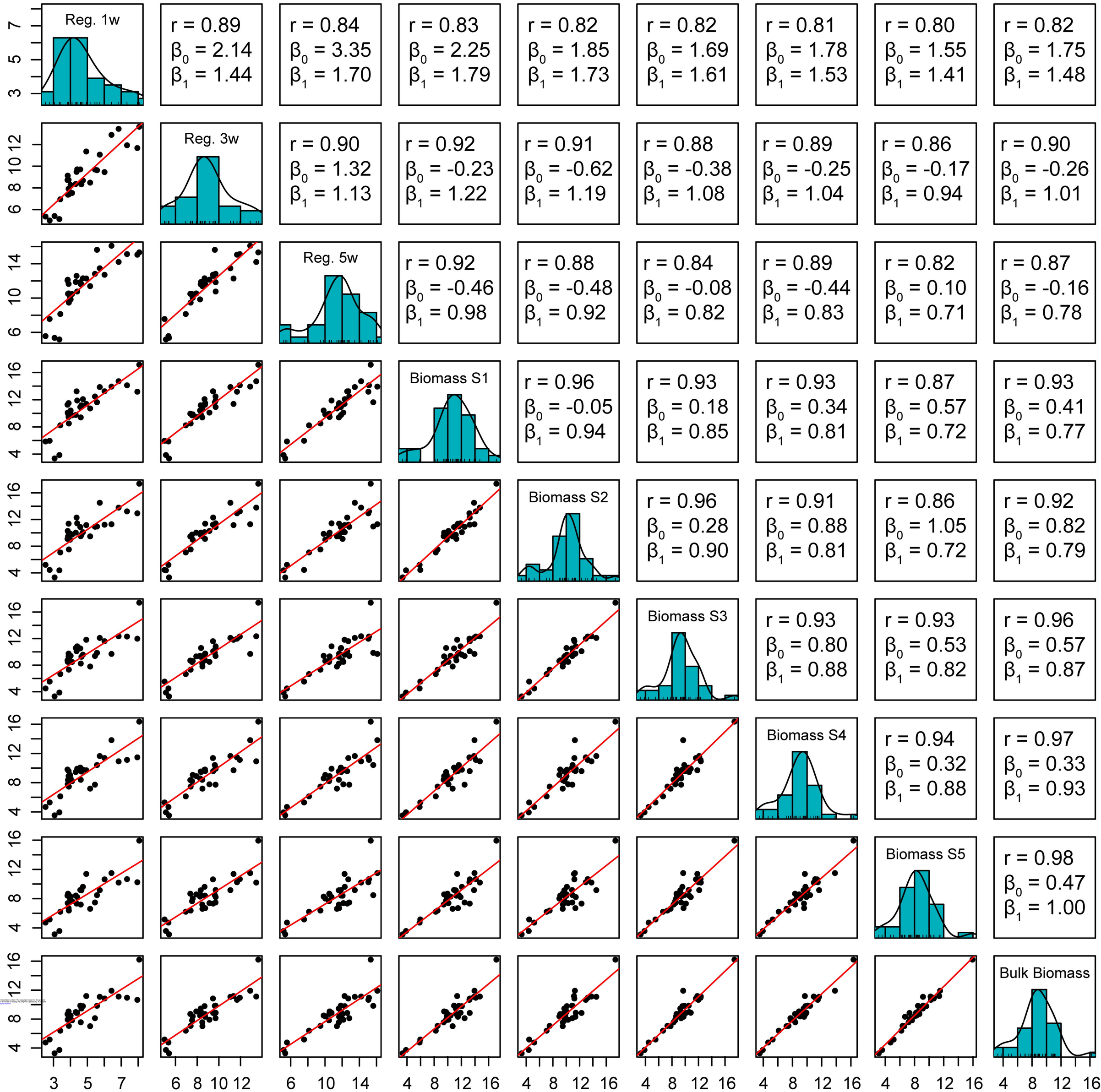
bioRxiv preprint doi: https://doi.org/10.1101/2022.12.02.518916; this version posted December 3, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

cultivar-level *k*-means clusters based on the proportions of stripped inflorescence biomass produced in each section. Regression lines and confidence intervals are cultivar-level generalized linear models.

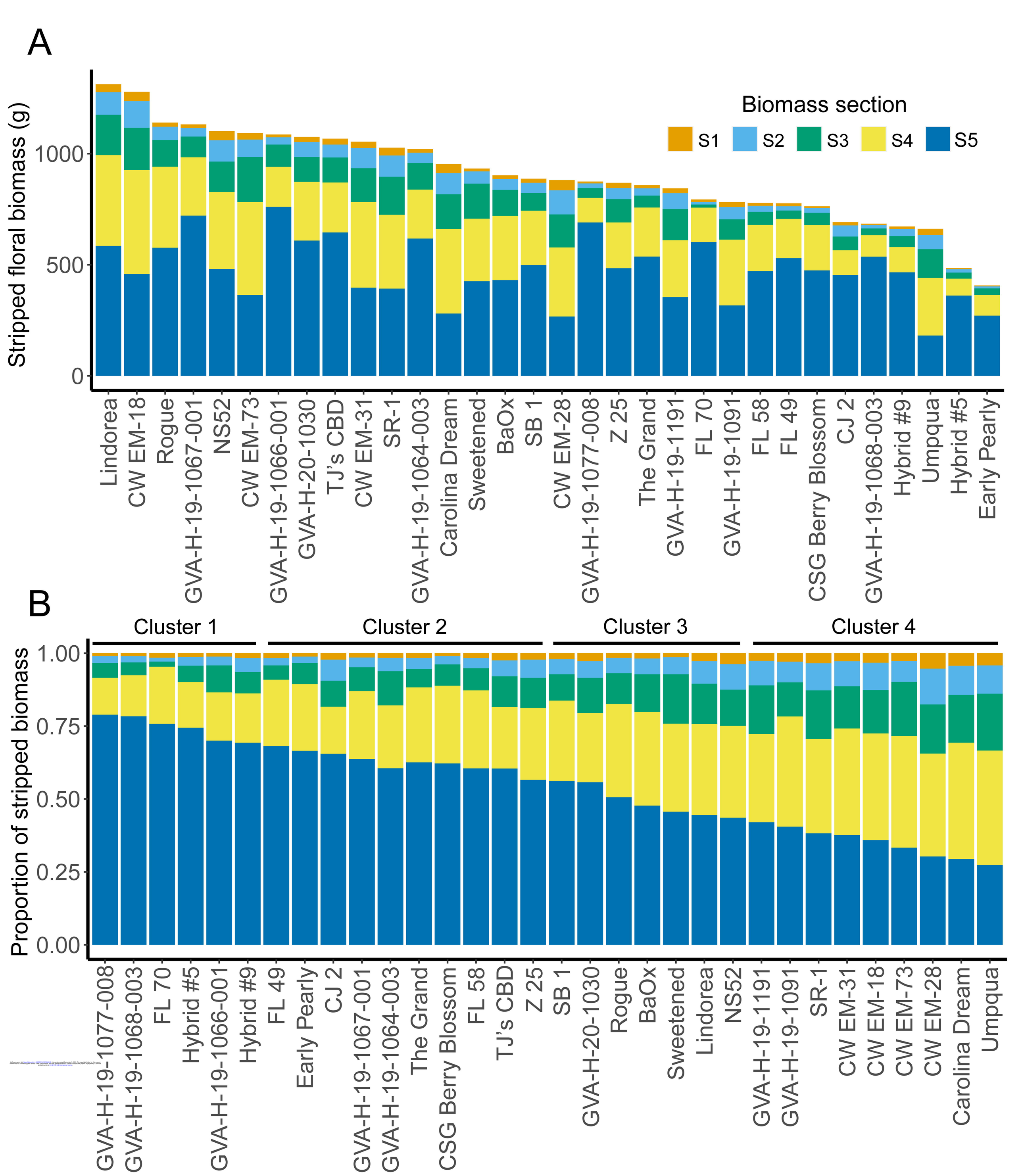
Figure 6. Boxplots showing the distribution of total potential cannabinoid concentration measured in post-harvest stripped biomass samples for plants partitioned into five sections where S1 represents branches derived from the top 5th of the main stem, S2 those from the second to top, etc. (A) Boxplot of all samples from all cultivars. (B) Thirty-two cultivar-level boxplots ordered by the slope of the regression predicting total potential cannabinoids using biomass section. *P*-values reported are for mixed models predicting total potential cannabinoid concentration where biomass section was a fixed effect and the plot that the sample was from was a random effect. Asterisks indicate a significant effect of biomass section after correcting for multiple testing and letters indicate statistically significant differences between sections of a cultivar based on a post-hoc Tukey's HSD test.

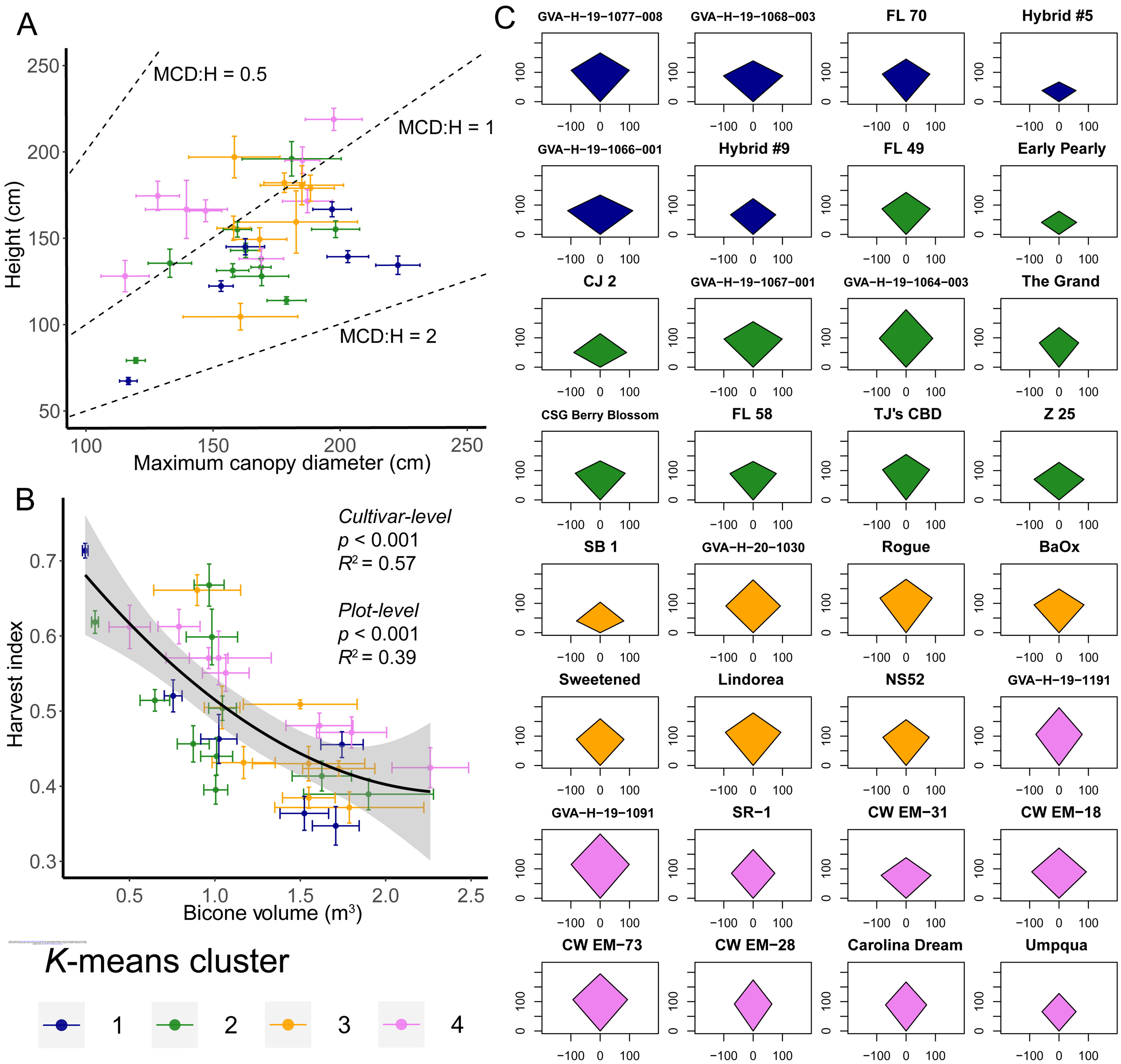
Figure 7. Regressions predicting changes in total cannabinoid concentration by section. (A) Multiple linear regression predicting cultivar-level values using canopy area, log of dry canopy density, and an interaction term as predictors. (B) Regression of plant-level observed values against predicted values for multiple linear regression where predictors were selected by stepwise regression. The solid black line represents a perfect correlation of predicted and observed values.

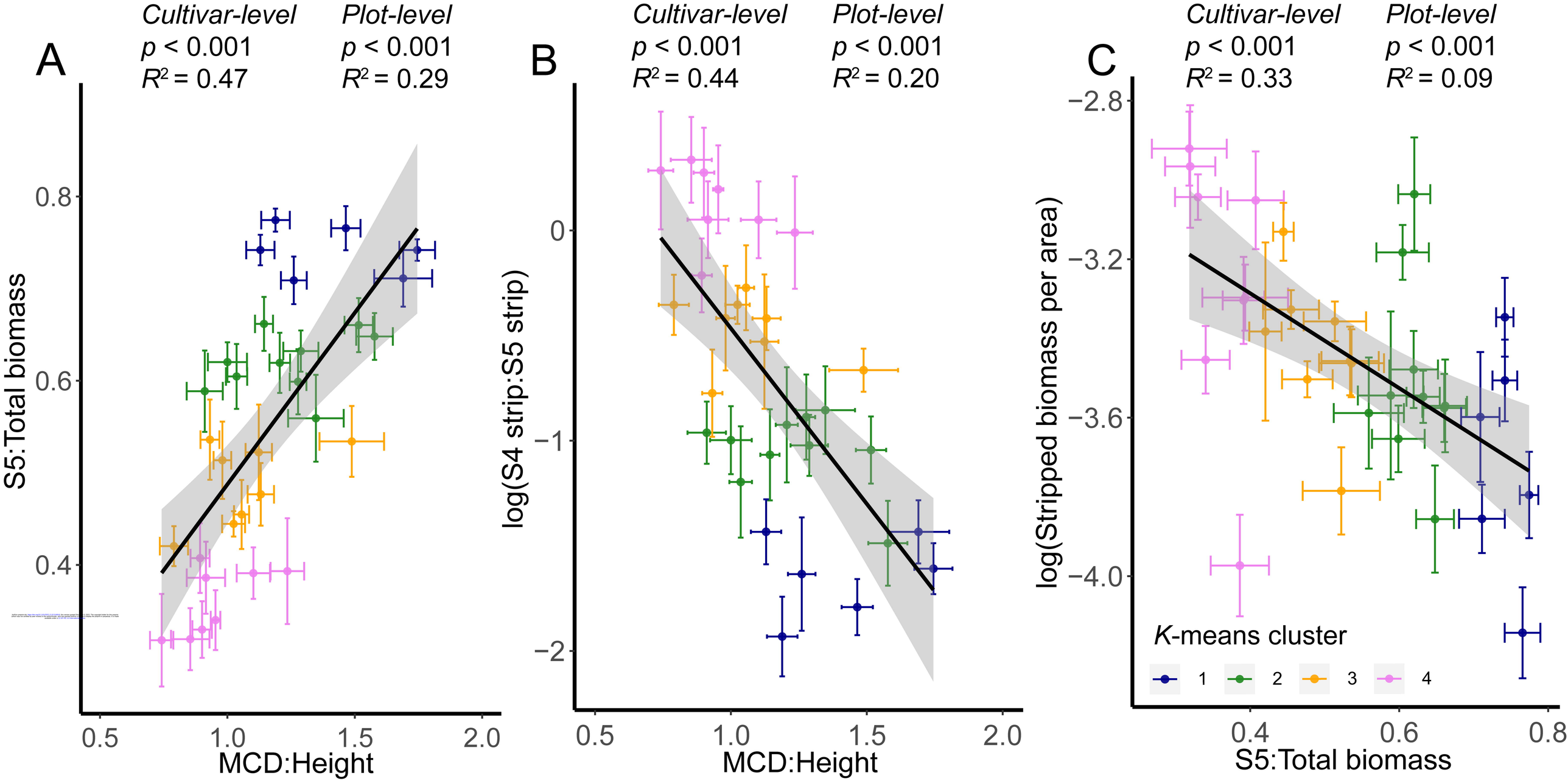

A
B
C
D
Height
MCD
MCDH
Section 1 (S1)
Section 2 (S2)
Section 3 (S3)
Section 4 (S4)
Section 5 (S5)

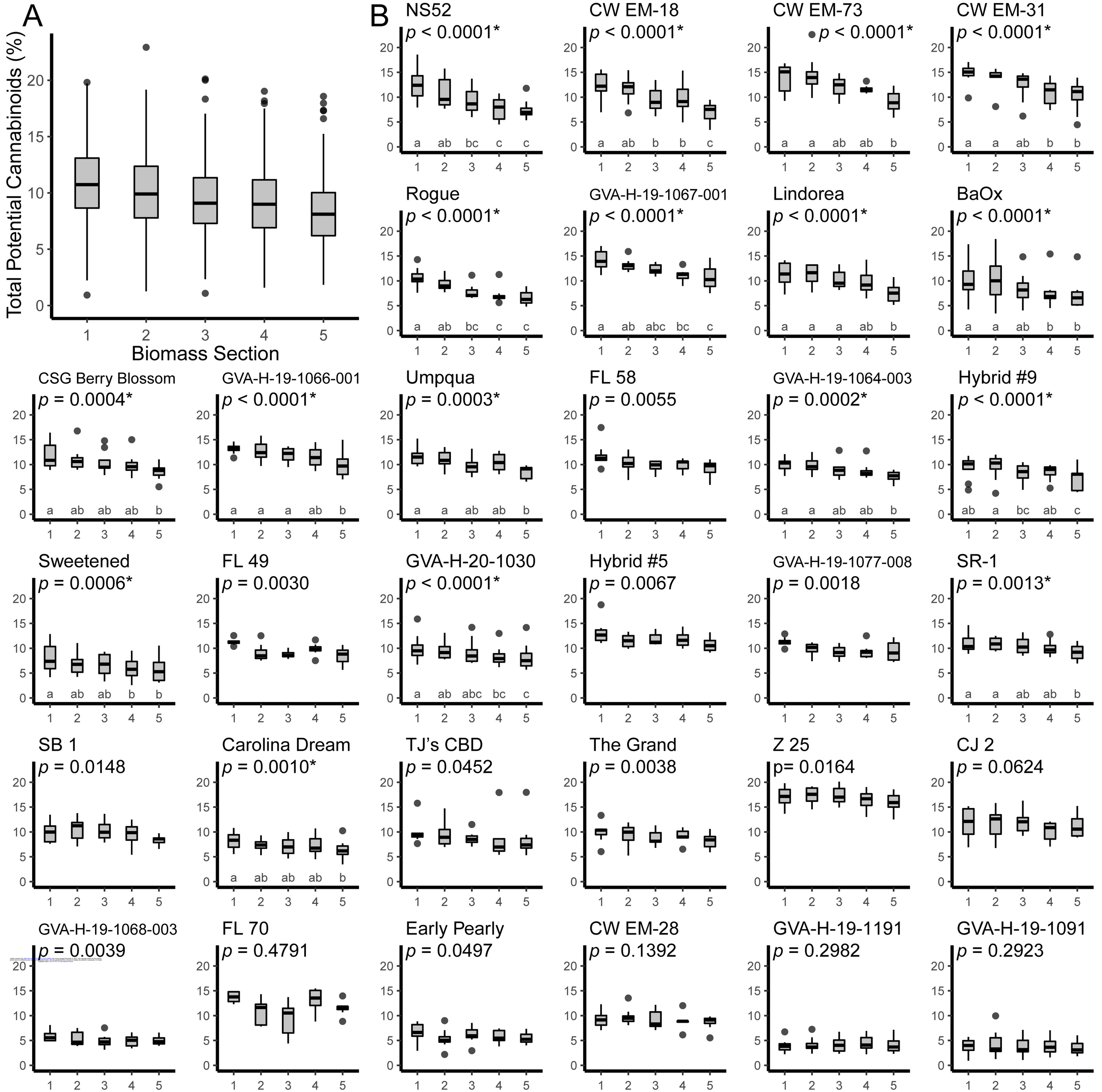


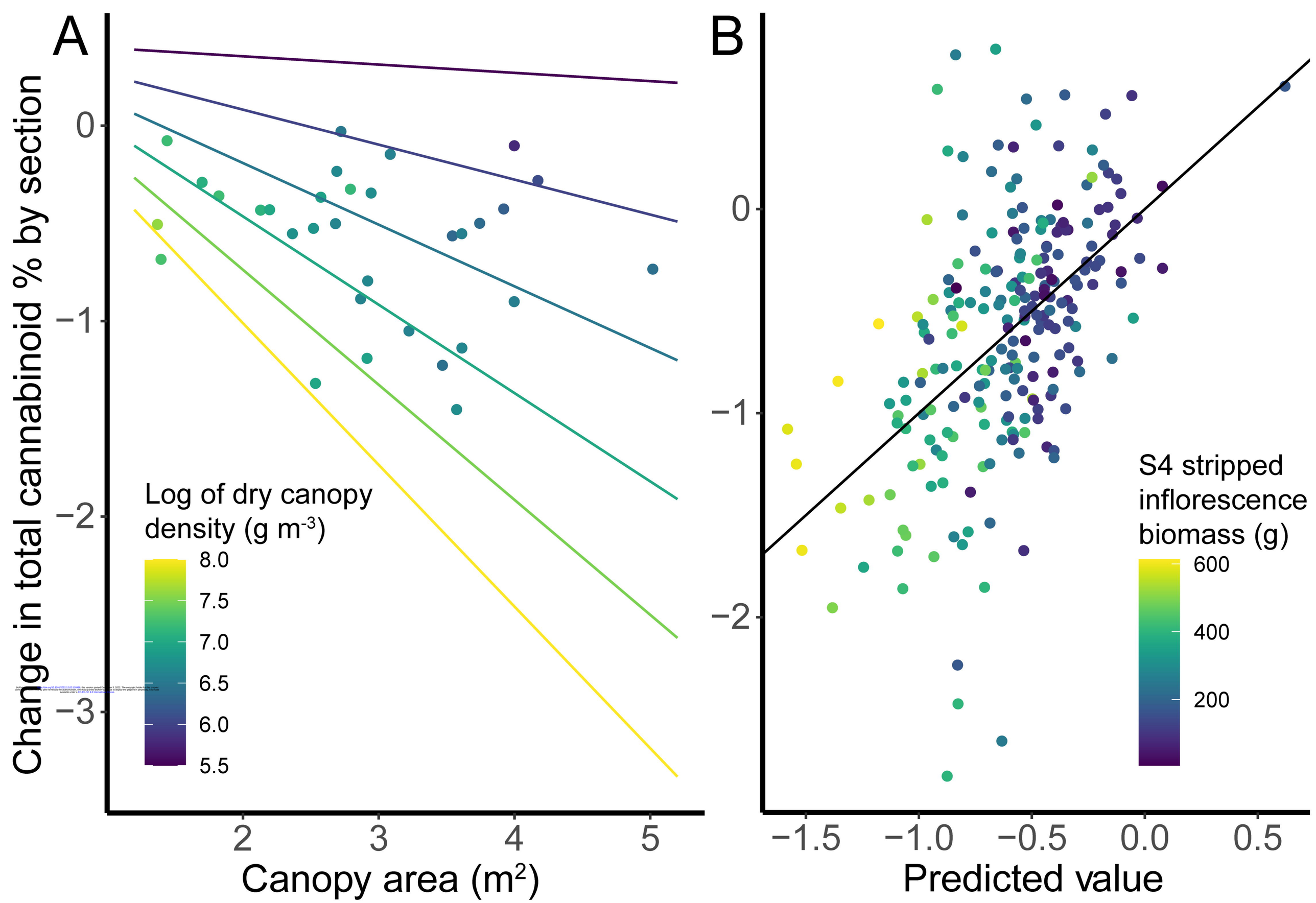
bioRxiv preprint doi: https://doi.org/10.1101/2023.12.22.573816; this version posted December 3, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.


A
Stripped floral biomass (g)
1000
500
0
Biomass section
S1
S2
S3
S4
S5
Lindorea
CW EM-18
Rogue
GVA-H-19-1067-001
NS52
CW EM-73
GVA-H-19-1066-001
GVA-H-20-1030
TJ's CBD
CW EM-31
SR-1
GVA-H-19-1064-003
Carolina Dream
Sweetened
BaOx
SB 1
CW EM-28
GVA-H-19-1077-008
Z 25
The Grand
GVA-H-19-1191
FL 70
GVA-H-19-1091
FL 58
FL 49
CSG Berry Blossom
CJ 2
GVA-H-19-1068-003
Hybrid #9
Umpqua
Hybrid #5
Early Pearly
B
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Proportion of stripped biomass
1.00
0.75
0.50
0.25
0.00
GVA-H-19-1077-008
GVA-H-19-1068-003
FL 70
Hybrid #5
GVA-H-19-1066-001
Hybrid #9
FL 49
Early Pearly
CJ 2
GVA-H-19-1067-001
GVA-H-19-1064-003
The Grand
CSG Berry Blossom
FL 58
TJ's CBD
Z 25
SB 1
GVA-H-20-1030
Rogue
BaOx
Sweetened
Lindorea
NS52
GVA-H-19-1191
GVA-H-19-1091
SR-1
CW EM-31
CW EM-18
CW EM-73
CW EM-28
Carolina Dream
Umpqua


A
MCD:H = 0.5
MCD:H = 1
MCD:H = 2
Height (cm)
Maximum canopy diameter (cm)
B
Cultivar-level
p < 0.001
R² = 0.57
Plot-level
p < 0.001
R² = 0.39
Harvest index
Bicone volume (m³)
K-means cluster
1
2
3
4
C
GVA−H−19−1077−008
GVA−H−19−1068−003
FL 70
Hybrid #5
GVA−H−19−1066−001
Hybrid #9
FL 49
Early Pearly
CJ 2
GVA−H−19−1067−001
GVA−H−19−1064−003
The Grand
CSG Berry Blossom
FL 58
TJ's CBD
Z 25
SB 1
GVA−H−20−1030
Rogue
BaOx
Sweetened
Lindorea
NS52
GVA−H−19−1191
GVA−H−19−1091
SR−1
CW EM−31
CW EM−18
CW EM−73
CW EM−28
Carolina Dream
Umpqua


A
Cultivar-level
p < 0.001
R² = 0.47
Plot-level
p < 0.001
R² = 0.29
S5:Total biomass
MCD:Height
B
Cultivar-level
p < 0.001
R² = 0.44
Plot-level
p < 0.001
R² = 0.20
log(S4 strip:S5 strip)
MCD:Height
C
Cultivar-level
p < 0.001
R² = 0.33
Plot-level
p < 0.001
R² = 0.09
log(Stripped biomass per area)
S5:Total biomass
K-means cluster
1
2
3
4


A
Total Potential Cannabinoids (%)
Biomass Section
B
NS52
p < 0.0001*
CW EM-18
p < 0.0001*
CW EM-73
p < 0.0001*
CW EM-31
p < 0.0001*
Rogue
p < 0.0001*
GVA-H-19-1067-001
p < 0.0001*
Lindorea
p < 0.0001*
BaOx
p < 0.0001*
CSG Berry Blossom
p = 0.0004*
GVA-H-19-1066-001
p < 0.0001*
Umpqua
p = 0.0003*
FL 58
p = 0.0055
GVA-H-19-1064-003
p = 0.0002*
Hybrid #9
p < 0.0001*
Sweetened
p = 0.0006*
FL 49
p = 0.0030
GVA-H-20-1030
p < 0.0001*
Hybrid #5
p = 0.0067
GVA-H-19-1077-008
p = 0.0018
SR-1
p = 0.0013*
SB 1
p = 0.0148
Carolina Dream
p = 0.0010*
TJ's CBD
p = 0.0452
The Grand
p = 0.0038
Z 25
p= 0.0164
CJ 2
p = 0.0624
GVA-H-19-1068-003
p = 0.0039
FL 70
p = 0.4791
Early Pearly
p = 0.0497
CW EM-28
p = 0.1392
GVA-H-19-1191
p = 0.2982
GVA-H-19-1091
p = 0.2923


A
B
Change in total cannabinoid % by section
0
−1
−2
−3
Log of dry canopy density (g m-3)
8.0
7.5
7.0
6.5
6.0
5.5
2
3
4
5
Canopy area (m2)
0
−1
−2
S4 stripped inflorescence biomass (g)
600
400
200
−1.5
−1.0
−0.5
0.0
0.5
Predicted value