scientific reports

OPEN

Transcriptomic analysis of CDL-gated photoperiodic flowering mechanisms in cannabis and their responsiveness to R: FR ratios in controlled environment agriculture

Samuel R. Haiden[1], Nathan Johnson[2] & Gerald A. Berkowitz[1✉]

While discrete regulatory mechanisms have been identified, a unified model for the transcriptional activation of photoperiodic flowering in short-day (SD) plants has not been delineated. Although the *GIGANTEA-CONSTANS-FLOWERING LOCUS T* (*GI-CO-FT*) pathway appears to be highly conserved, each species may differ in its blueprint for activation of *FT* and floral meristem identity. In cannabis, an SD plant, *PSEUDO-RESPONSE-REGULATOR 37 (PRR37)* has a causal effect in determining photoperiod sensitivity. This study identifies a network of *WITH NO LYSINE (K)* kinases (*WNK*) which may be regulating the PRR proteins and downstream genes, including florigen (*FT*). CO-like genes have been identified in cannabis, but their function in regulating *FT* has not been determined. This study shows that in cannabis, CO is critically-day-length-gated (CDL) and down-regulated by short days, likely to act as a repressor of *FT*. We demonstrate that CDL-gated gene expression can be manipulated in cannabis horticulture through modifying the R:FR ratio of LED lighting, where increased expression of *COL5* in response to light treatment is inversely related with expression of *FT*. Building on these insights, we propose a model for the CDL-gated regulation of *FT* expression in cannabis.

Phytochrome, PSS, and photoperiod recognition

Plants recognize sunlight via specialized molecules in the cytosol called photoreceptors[1]. There are several classes of photoreceptors including phytochrome, cryptochrome, phototropin, *ZEITLUPE* (*ZTL*) and *FLAVIN BINDING KELCH REPEAT F-BOX 1* (*FKF1*). These molecules control genes which determine the photoperiodic flowering response, circadian rhythms, photosynthesis, photorespiration, cellular respiration at night, and more[2,3].

When the phytochrome chromophore is irradiated by red light, it moves into its active conformer, which interacts with phytochrome interacting factors (PIFs) to modify gene expression[3]. Far-red light, however, rapidly converts the active conformer Pfr to the inactive conformer Pr; darkness also converts Pfr to Pr, but more slowly[4]. During the day, most but not all the phytochrome is activated, and the ratio of Pr to Pfr will dictate certain physiological responses, including the shade avoidance response, which is induced due to the nature of shady light to contain a high ratio of far-red light[5]. This is because far-red light, having a longer wavelength, can penetrate more deeply through the canopy. The ratio of Pr to Pfr is known as the PSS, or Phytochrome Stationary State[6]. It has been shown that the PSS can be predicted based on the ratio of red to far-red light that is being received by a plant tissue[7]. Thus, we can model PSS in a leaf tissue given the following formula:

$$PSS = \frac{\Sigma\left(\frac{800}{300}\right)N\left(\lambda\right)\sigma_r\left(\lambda\right)}{\Sigma\left(\frac{800}{300}\right)N\left(\lambda\right)\sigma_r\left(\lambda\right) + \Sigma\left(\frac{800}{300}\right)N\left(\lambda\right)\sigma_{fr^\lambda}}$$

Or, written more simply,

[1]University of Connecticut Agricultural Biotechnology Laboratory, Storrs, CT, USA. [2]Verne Bioanalytics Inc., Auburn, MA, USA. ✉email: gerald.berkowitz@uconn.edu

$$PSS = (\frac{P_r}{P_r + P_{fr}})$$

The entrainment of photoperiod pathway genes through uninterrupted skotoperiod length is the primary signal for the initiation of photoperiodic flowering, which takes effect through increased expression of FT in the leaves[8]. Cannabis is a highly polymorphic single species; there are subspecies of cannabis that are photoperiod insensitive, and others which flower under short days: however, no cannabis plants flower exclusively under long days[9,10]. Cannabis is a short-day photoperiodic flowering plant and although sexually mature flowers form prior, it produces its inflorescences when the critical day length is under 12–14 hours[10]. Photoperiod sensitive cannabis tends to flower when the night length exceeds ten hours. Although we refer to critical *day* length (CDL) instead of critical *night* length, It is important to understand that photoperiodic flowering depends not on the constancy of the photoperiod, but on the constancy of the dark period[11]. The critical day length at which cannabis begins to flower is a key determinant in success when it comes to cultivation, because the longer the CDL, the sooner in the season flowers emerge[12], which guarantees that the flower can ripen before cold weather in the northern latitudes. Therefore, understanding the genetic mechanisms which trigger CDL-gated flowering activation cascades is critical for cannabis flower production.

The FT activation cascade in cannabis depends on the protein–protein interactions of CCT-domain bifunctional regulators

The signaling cascade which initiates the photoperiodic flowering response is well characterized in long day (LD) plants[13]. In Arabidopsis, a model LD plant, CO acts as a central hub for integrating environmental signals to determine flowering time[14]. The *FT* gene is activated in leaves via the sequential activation of GI and CO[15], which are regulated by both phytochromes and cryptochromes[16]. Upon entering long days, the FT protein is positively regulated and translocated to the meristem, where it works with FD and SUPPRESSOR OF CONSTANS 1 (SOC1) to activate floral meristem identity, which initiates the development of flowers[13,17].

The CCT (CONSTANS, CONSTANS-like, TOC1) domain or CONSTANS domain is a highly conserved genetic feature common to CONSTANS genes, among others. The encoded domain enables protein–protein interactions with TFs like Nuclear Factor Y (NFY) that directly bind the CCAAT box in the *FT* promoter[18]. Although it has been shown that in Arabidopsis the CCT domain of the CONSTANS protein is responsible for FT activation, pseudo-response regulators like those in rice and cannabis possess CCT domains as well[19]. This group of pseudo-response regulators have been shown to be substrates for phosphorylation by WITH NO LYSINE (K) (WNK) kinases, which are regulated by circadian rhythms[20].

Like Arabidopsis, rice (an SD plant) possesses a CO-like gene, but this gene is known as *HEADING DATE 1 (HD1)*[21,22]. Unlike in Arabidopsis, *HD1* in rice is not a central hub for signal integration, but rather a bifunctional regulator which performs different functions based on the presence of multiple interactor proteins[22]. Soybean is also an SD plant, and has a system of activating FT expression based on CDL gated genes and circadian rhythm entrainment[23]. A group of "E genes" and *GmELF3* regulate the photoperiodic flowering response in soybean. The *E2* gene is an ortholog of *GI*. Not dissimilarly to rice, *GI* regulates the expression of *FT* in soybean, inhibiting its expression during long days[24]. Although several models have been proposed to explain the protein–protein and protein-DNA interactions which control SD flower initiation, a conclusive model has not been established[21]. The known mechanisms of SD flowering initiation in rice and soy provide clear insight into the potential mechanisms for initiation of flowering in cannabis.

Cannabis CONSTANS-like (CO, COL, CO-like) genes have been identified, and there are at least 13 expressed CONSTANS-like genes[25]. However, no evidence has demonstrated that CO is required for *FT* activation. In cannabis, florigen-like, *PRR37*-like, and *SOC1*-like genes have also been identified[19,26], and *CsPRR37* has been shown to be causal to and required for the activation of *FT*, whether or not it requires the presence of CO. It has been shown that a truncation of the CsPRR37 protein, removing the CCT domain, is correlated with the autoflowering trait, or lack of response to photoperiod[19].

COL genes have varying expression levels in different tissues. Seven (4, 8, 1, 6, 11, 9, 12) are related and highly expressed in leaves, three are related and expressed in leaves and flowers (7, 5, 10), one is highly expressed in stems (13), and two are highly expressed in flowers(2, 3)[25]. *COL3* and *COL7* exhibit amino acid differences among early flowering and late flowering cultivars[25]. It has been shown that all the cannabis CO genes have diurnal expression patterns in both LD and SD conditions[25], however no data showing the expression of these genes during SD transition has been published. There seems to be a dichotomy in CO genes regarding expression levels in early and late-flowering genotypes, where *CsCOL4* and *11* expression levels are higher in early flowering genotypes, and *CsCOL6*, *7*, *9* and *12* are lower. "This indicates that there may be multiple CsCOL genes functioning as promoters or suppressors of flowering to regulate flowering time in C. sativa"[27].

Cannabis *FT* (florigen) has been identified, and sequence variations in the FT protein have been associated with the photoperiod sensitive/insensitive phenotype[26]. There are 4 known *FT*-like genes in the cannabis genome, which are part of a larger family of PHOSPHATYIDYLETHANOLAMINE BINDING PROTEIN (PEBP) genes involved in flowering. These include three closely related to *MOTHER OF FT (MFT),* two related to *TERMINAL FLOWER (TFL)*, and three *BROTHER OF FT (BFT)/CEN* genes. In solitary flowers in vegetative plants, six PEBP genes were differentially expressed. In the reproductive nodes, the three *CsBFT/CEN*, *CsTFL* and *CsMFT3* genes exhibited reduced expression, while the *CsFT4* gene expression increased[27,28].

Although the *GI-CO-FT* photoperiod-dependent flowering pathway appears to be highly conserved across different crops including cannabis[24,29], the mechanisms and functions of each component of the signaling pathway can vary. Understanding the specific mechanisms of this pathway is an ideal path for elucidating the regulation of flowering time in cannabis[27].

MicroRNA and the age dependent pathway

Although the inductive, short photoperiod "flowering" phase is when the inflorescence forms, and is also when the majority of flowers are produced, it is not the point of sexual maturity[30]. The cannabis plant produces solitary flowers at the leaf and stem axis after reaching the mature vegetative phase (non-inductive), and these flowers are in fact sexually mature and capable of producing seed. These solitary flowers can emerge at the 4th–6th internodes (Cervantes), and transcriptomic results have shown that several floral induction genes are upregulated, including *MFT*, *SOC1*, *LEAFY* (*LFY*), and *APETALA* (*AP1*), while several repressors *TEMPRANILLO* (*TEM*), *TFL1*, and *BFT* were downregulated[28]. *SPL* genes were also activated.

The age-dependent flowering pathway in cannabis has been elucidated[28] and mechanisms homologous to Arabidopsis have been identified, including miRNA[31]. The function of miRNA172 and 156 in regulating flowering genes is one of the most highly conserved mechanisms in the plant kingdom. As such, it is not unreasonable to assume that these miRNAs are having a significant impact on cannabis flowering.

One important mechanism for age-dependent flowering in Arabidopsis are "count-down" and "count-up" miRNA timers. miRNA156/157 act as "count-down" timers, meaning their expression decreases as the plant ages, and miRNA172 expression increases as the plant ages. When the expression levels of these genes reach critical thresholds, they initiate flowering[32,33]. Target genes for miRNA156/157 include *SPL* genes, which are critical in regulating the transition from the juvenile to the mature phase[32,34]. miRNA172 affects the AP2-family TF's, which are also flowering repressors. Therefore, the increased expression of miRNA172 results in the decreased expression of these flowering repressors, which initiates flowering[35,36].

Other miRNA's have been shown to function as flowering mechanisms in rice, where differentially expressed osa-miR171 and osa-miR1432 were associated with effects on yield[37]. A group of microRNA's including csa-miR156 and csa-miR172 have been identified in the cannabis genome, and csa-miR156 has been shown to target SPL genes while csa-miR172 targets transcription factors that regulate flowering time[31].

Results/ discussion

Transcriptome mapping and differential expression analysis

The differential expression analysis below displays the fold change in gene expression of several different photoperiodic-flowering related genes in cannabis leaves. The first sample set represents LD conditions (last long day), while the second sample set represents SD conditions (seventh short day). Thus, these data demonstrate gene expression changes in cannabis leaves across SD transition. Multiple known genes were identified, and their CDL-gated response characterized, while several other genes were uncovered, homologous to genes in soy and Arabidopsis.

The sequenced transcriptomes were mapped to nine different reference genome assemblies due to what transcripts are defined directly affects what alignment and quantification is possible with short read- RNA-Seq. After aligning the sequenced reads using Salmon, the mapping rate ranged from 20 to 70% with cs10 achieving the lowest mapping rate at 25% (supplemental Fig. 1). Supplemental Fig. 2 displays an alternative volcano plot for the analysis. Supplemental Figs. 3 and 4 display the effect of shrinkage. The highest mapping rate was achieved by Jamaican Lion, Purple Kush, and CBRX18 at ~ 70%. Out of those three transcriptomes, Jamaican Lion[38] had the highest number of annotated transcripts, which we continued to use for the study. To identify significant differentially expressed genes (DEGs), we used DESeq to identify thousands of relevant targets with a wide range of annotation on function.

Below are data on photoperiodic flowering genes (Fig. 1), but dozens of HD-ZIP, MYB, RAV, ABC, and CESA genes were also significantly differentially expressed. For table displaying each DEG in the scatterplot below with their associated base mean, adjusted P-value, log2 fold change, and associated NCBI accession, see supplemental table 1.

### The CO gene family in cannabis is CDL-gated, downregulated upon SD transition

The function of CO-like genes in SD plants have been characterized in rice (*HD1*)[22], but their function are not clear in other plants including cannabis. In rice, *HD1* expression tends to remain stable while its activity is modulated by protein–protein interactor. Uncharacteristically, in this study, CO genes tend to be downregulated by SD conditions, with *CsCOL10* being downregulated with a log2 fold change of − 4. The Predictive models in this paper (supplemental Figs. 5 and 6) and literature review above suggest that CsCOL genes are regulated by phytochrome via *HY5*, and possibly co-regulate *FT* with the PRR genes. *CsCOL5* and *10* domains are depicted in supplemental Figs. 9 and 10.

PRR37/ APRR3/ WNK Kinase phosphorylation pathway is CDL-gated

The cannabis APRR*3* is homologous to the Arabidopsis APRR3, which is phosphorylated by WNK kinases[20]. The *CsPRR37* and *APRR3* genes from this transcriptome both contain CCT domains (supplemental Figs. 7 and 8). While *PRR37* was downregulated at -3.33 log2 fold change upon entry into short days, *APRR3* was only downregulated by -0.44. The downregulation of *APRR3* is contested by our qPCR data and is yet to be validated. According to the DESeq, *WNK4* and *11* were also downregulated with a log2 fold change greater than 2, while *WNK9* was down regulated by -0.45. While the downregulation of *WNK9* was confirmed, the downregulation of *WNK4* was not. The function of WNK kinases and their sensitivity to photoperiod in cannabis have not been described, despite extensive study of *PRR37*.

TOE1

In Arabidopsis, *TOE1* binds to the activation domain of CO-like genes. During long days, it interacts with *CRY* and releases its repression of CO[39]. Different genes regulate CO in a similar way in rice[40]. The function
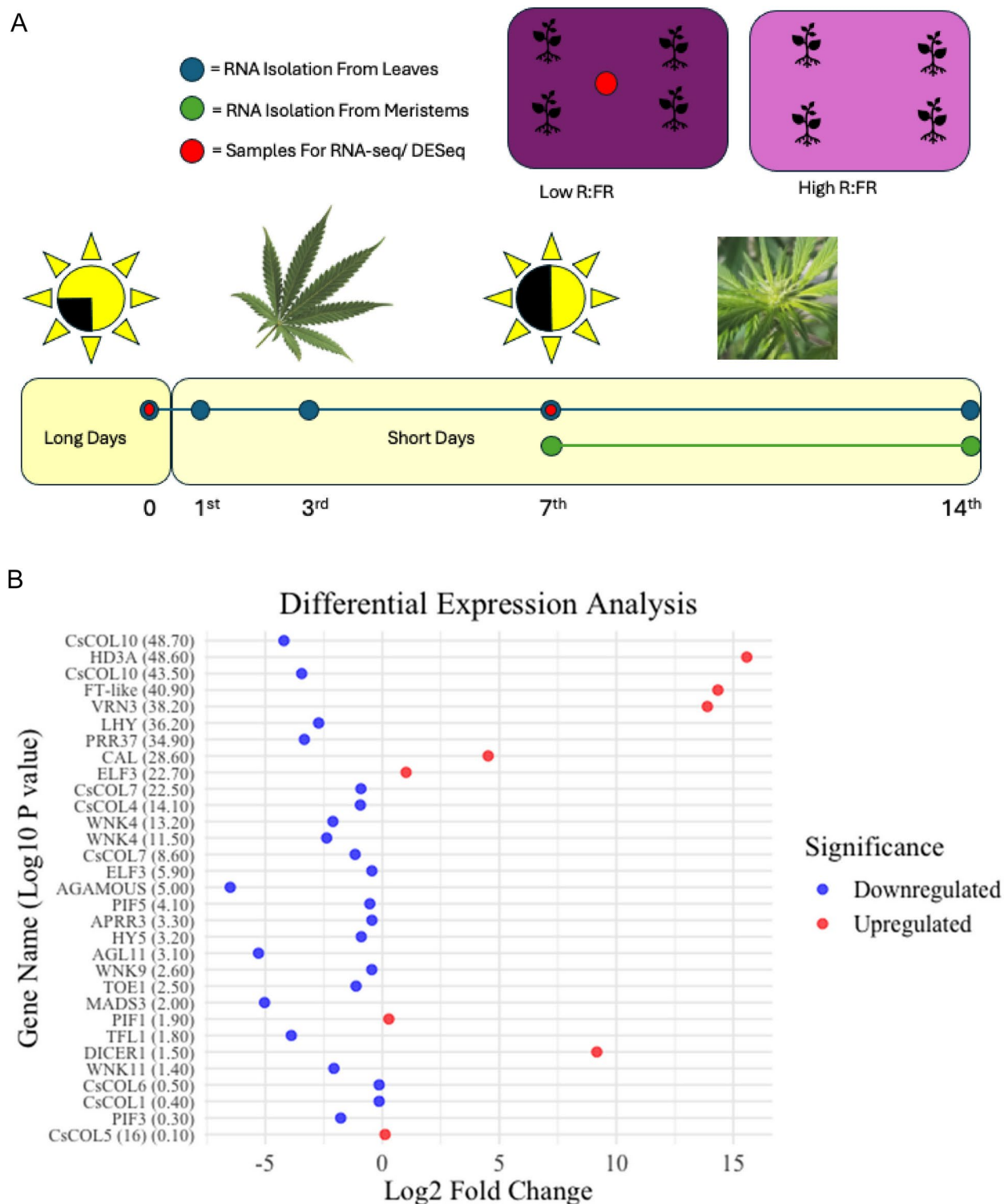
Fig. 1. Differential Expression Analysis. Panel (**A**): Experimental design. Panel (**B**): Scatterplot: red dots indicate up regulated genes, while blue dots indicate down-regulated genes. The X-axis indicates the log2 fold change, either negative or positive. The Y-axis indicates the log10p-values associated with each gene's fold change. Greater values indicate higher confidence. CsCOL10 was downregulated with a high level of confidence, while HD3A, FT and VRN are upregulated with a high level of confidence.

of *TOE1* and its potential involvement in the photoperiodic flowering response in cannabis has not yet been characterized. Here we can see that *TOE1* is significantly downregulated upon entry into short days. *TOE1* is an *AP2*-like gene, which are known to be regulated by miRNA in the transition to mature/ flowering phase[28,35]. It is also located at the Autoflower1 locus[41].

TFL1, FT-like, HD3A-like, and Vernalization 3-like are CDL-gated PEBP mobile signals

In response to changes in CDL-gated repressors, three of these *FT*-like genes increase in expression with an enormous log2 fold change. Although *FT*-like genes, including *HD3A* and *VERNALIZATION 3*, are likely regulated by many conserved pathways involving *ELF3*, *GI*, CO, and *PRR37*- the regulation and specific function of the *TFL1* gene is unknown.

ELF3, LHY

Two *ELF3* transcripts were significantly differentially expressed, both homologous to the soybean *ELF3* gene, which is known to regulate circadian rhythms and the CO-*FT* pathway via *GI* entrainment[42]. Interestingly, one transcript was upregulated while the other was slightly down regulated. Similarly, *LHY* is known to regulate circadian rhythms, and was significantly down regulated.

DICER

Interestingly, next to the *FT*-like genes, the most upregulated gene in the leaf transcriptome upon the transition into short days was *DICER*. *DICER* is involved in the generation of siRNA from miRNA genes expressed in plant cells. In this way, several genes which regulate the transition from juvenile to mature phase, as well as flowering, are regulated[31]. Two known miRNAs involved in gene silencing in cannabis are miRNA156 and 172, which target AP2 and other SPL genes[28,31]. The generation of siRNA has also been shown to be necessary for the formation of heterochromatin, and therefore *DICER* may also be playing a role in restructuring chromatin architecture during floral transition[43].

qPCR validation of DESeq results

qPCR was performed on the same samples as the DESeq analysis with the addition of the third short day (Fig. 2). Most results were consistent with the DESeq. Time points represented are "Last Long Day" (LLD), "Third Short Day" (SD3), and "Seventh Short Day" (SD7). A one-way ANOVA and Tukey-HSD test comparing all CO-like expression values for LLD with SD7 revealed that the CO-like gene family is significantly down-regulated by SD7 ($P < 0.001$).

Although our DESeq data and qPCR data are consistent for the CO gene family, showing a trend of down-regulation by SD transition, other genes yielded inconsistent results, including *APRR3* and *WNK4*. This could be the result of cDNA degradation which occurred over the passage of time between the *CO/FT/PRR37* qPCR analyses and those on *APRR3, PIF, HY5, TOE1, DICER* and *WNK*.

Protein–Protein interaction network prediction

In the above protein–protein physical interaction prediction (Fig. 3), blue dots indicate serine/threonine kinases (WNK), green dots indicate pseudo-response regulator genes including PRR37 and APRR3, and the red dots indicate phytochrome and PIF3. Predictions suggest that PIF3 physically interacts with the pseudo-response regulators, which initiates a WNK signaling cascade. The nature of the physical interaction is unknown; however, it seems likely that APRR3 and PRR37 are substrates for phosphorylation by WNK kinases. For all physical interaction and functional relationship networks, the thickness of the line connecting the dots indicates the strength of the probability of the interaction/ relationship. A functional analysis including PRR37 and CO-like genes suggests that the two genes may be co-regulated by similar pathways (supplemental Fig. 5) while another protein–protein interaction analysis suggests that CO-like genes may be regulated by HY5, which is downstream of PHY (supplemental Fig. 6).

LED programming technology allows for specifically dialed "spectral diet"
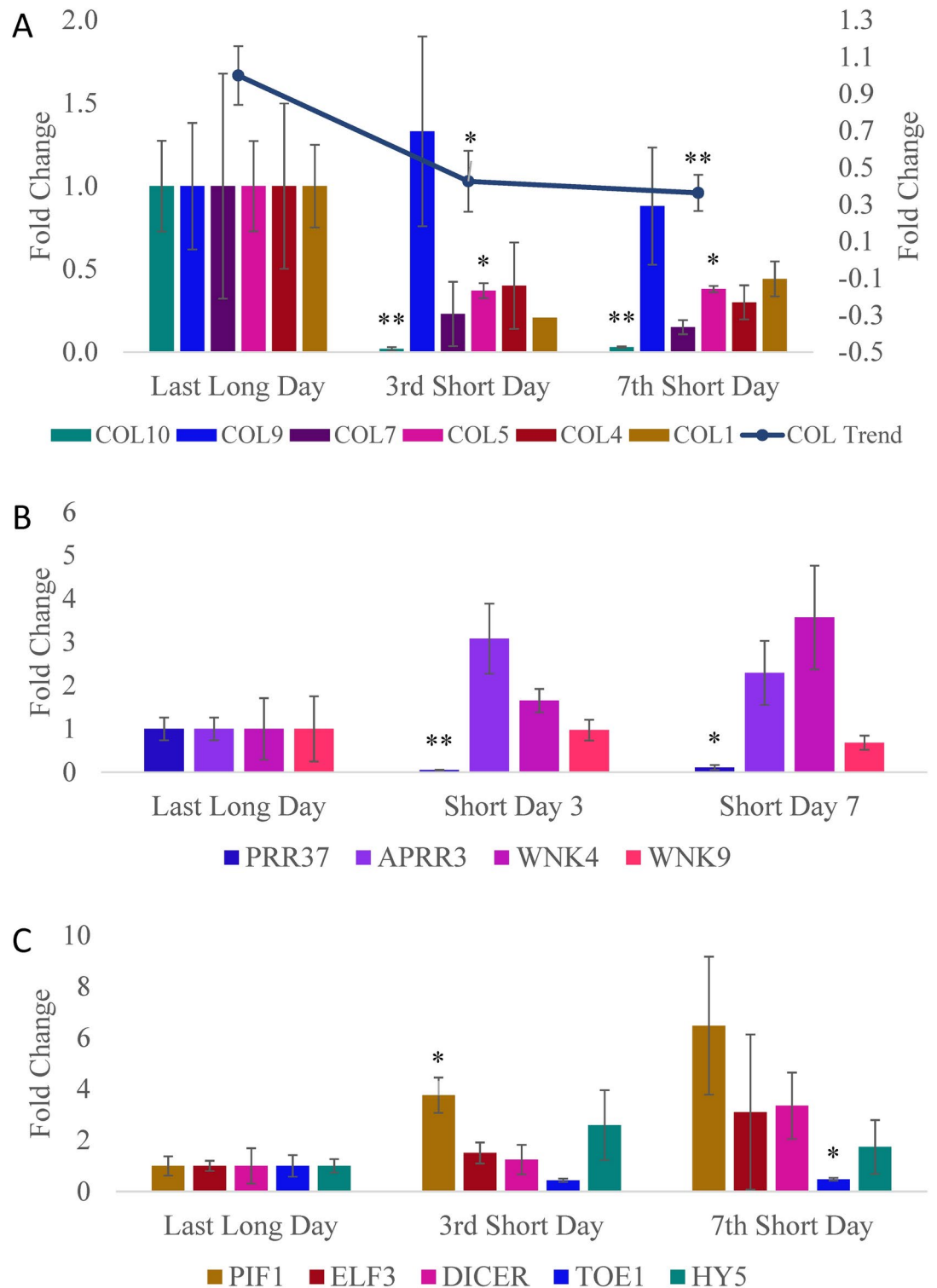
Due to the evidence suggesting that the molecular activation of FT is downstream of PHY, we set out to see if we could manipulate R:FR ratios and therefore phytochrome stationary state in our horticultural conditions to influence gene expression, phenotype, or chemotype. Per Sager et. al., the ratio of red (660 nm) to far red (700 + nm) light being applied to a given plant tissue can be used to approximate phytochrome stationary state PSS within that tissue[7]. The ratio of active to inactive phytochrome (PSS) was predicted per the equation below. Adaptiiv LED lights were used to dial in specific R:FR wavelengths. Treatments were normalized with regards to PPFD, and while the ratio of red to far red changed, the total amount of red light was controlled. Treatment was applied as constant condition for the duration of plant growth. Treatment conditions and variability are displayed below (Fig. 4).

$$PSS\ High\ RFR = \frac{7}{0.2 + 7} = 0.97 \quad PSS\ Low\ RFR = \frac{5.1}{1.9 + 5.1} = 0.72$$

Red to far-red ratio affects expression levels of CDL-gated photoperiodic flowering response genes FT, COL5, and WNK9

The high R:FR treatment ($\sim 19{:}1$) had some effects on gene expression compared to the control ($\sim 3{:}1$). By the seventh short day (SD7), *CsCOL5* was significantly upregulated in the high R:FR group, corresponding with a decreased expression level of *FT*. This suggests a relationship between *CO* and *FT*, where *CO* may act as

A — Fold Change comparison across Last Long Day, 3rd Short Day, 7th Short Day for COL10, COL9, COL7, COL5, COL4, COL1, and COL Trend.

B — Fold Change across Last Long Day, Short Day 3, Short Day 7 for PRR37, APRR3, WNK4, WNK9.

C — Fold Change across Last Long Day, 3rd Short Day, 7th Short Day for PIF1, ELF3, DICER, TOE1, HY5.

a repressor of *FT* until its co-regulation with *PRR37*, which results in the inhibition of the repression of *FT*. Interestingly, *WNK9* was also affected by increased red light, implying that it is under the control of *PHY* in cannabis, further supporting its role in integrating photoperiod signals. Due to the PSS calculation derived from the R:FR ratio, we theorize that the changes associated with increased R:FR ratio are due to increased phytochrome activity.

This method demonstrates that it is possible to affect the expression of signaling pathways downstream of phytochrome through the manipulation of horticultural lighting environments (Fig. 5).

Although we were able to elicit a change in gene expression in leaves, we were unable to ascertain a measurable impact on *SOC1*, a master transcriptional regulator of floral meristem identity, by using R:FR ratios. However, we did note that *SOC1* expression was significantly increased in the developing floral meristem between the 7th (SD7) and the 14th (SD14) short day, as displayed below.

◂**Fig. 2**. qPCR data describing the fold change in gene expression in cannabis leaves across SD transition. Four biological replicates per assay, from low RFR group. UB used as reference, delta-delta Ct method used to calculate relative expression change. In each assay, the last long day is set as control (1). For all experiments, student's t-test was utilized for calculating statistical significance of means separation. One star (*) indicates a P-value below 0.05, while two stars (**) indicates a P-value below 0.005. Panel A, CO-like genes: All but one CO-like genes were downregulated promptly by the third short day, which persisted into the seventh short day. COL9 expression remained relatively unchanged, which is characteristic of AtCO. The greatest fold change was exhibited by CsCOL10. Trend line indicates average fold change across all COL genes. Individual expression values across all biological replicates and genes used to calculate fold change, standard error and significance. Overall, the CO-like gene family in cannabis was down-regulated significantly upon SD transition. A one-way ANOVA comparing LLD to SD7 revealed significant differences in gene expression between groups across CO-like genes ($F(1, 44) = 14.81$, $p < 0.001$). Post-hoc Tukey's HSD test indicated significantly lower expression in SD7 compared to LLD (***$p < 0.001$). Second Y axis quantifies trend line. Panel B, PRR and WNK: PRR37 is heavily downregulated upon SD transition. Contrary to the DESeq results, qPCR results for APRR3 suggest that it was upregulated slightly by the third short day, while WNK4 was upregulated slightly by the seventh day. The qPCR results for WNK9 are consistent with the DESeq. Panel C: Most genes are marginally upregulated while TOE1 is downregulated. Although the fold change increase in expression of DICER is significantly lower in the qPCR assay than the DESeq, we still consider the data to confirm that expression levels are increased by the seventh short day. Four biological replicates per assay, from low RFR group, except in the case of DICER and TOE1 where 8 replicates are used from both the high and low RFR groups.
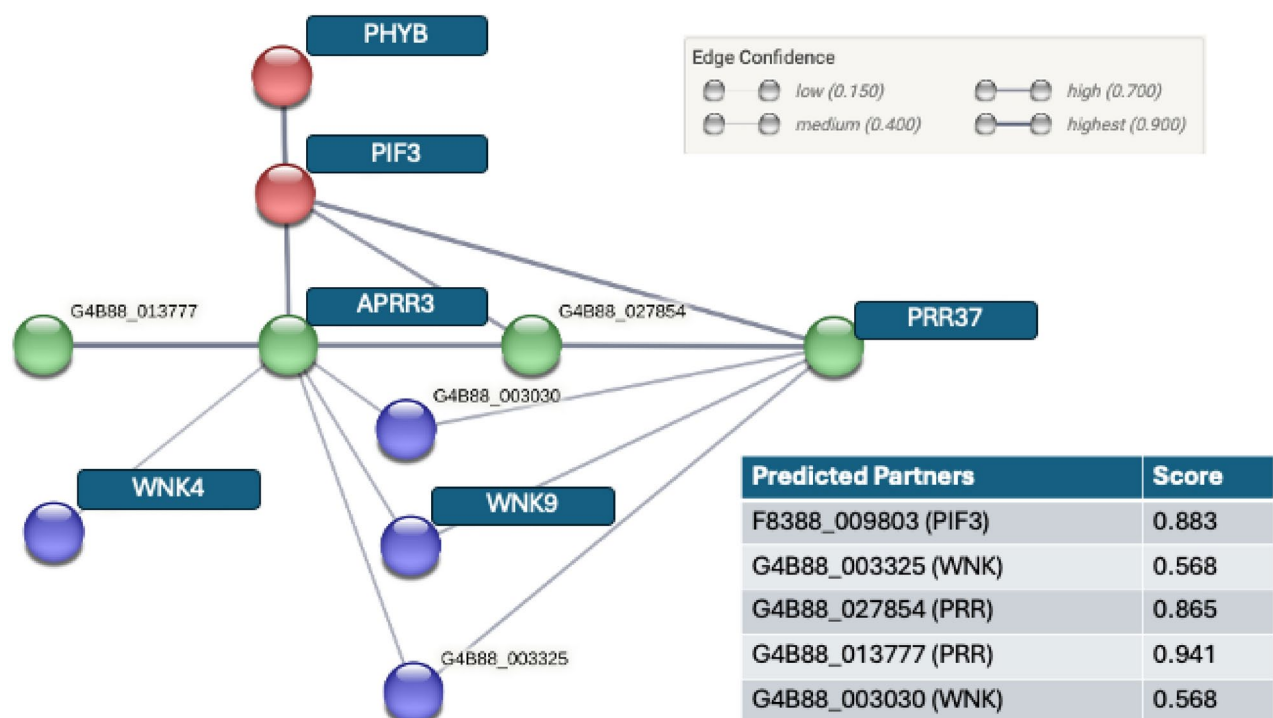

Fig. 3. Predictions generated through STRING database. Proteome constructed using Jamaican Lion reference, STRING proteome ID: STRG0A70MMF. Sequences for WNK4, WNK9, APRR3 and PHYB were extracted from DESeq transcriptome. PIF3 predicted by STRING and identified as homologous to phytochrome interacting factor. Other partners predicted include two pseudo-response regulators and two WNK kinases. Predicted protein–protein interactions indicated by gray line. Edge confidence categorized as low, medium, high, or highest, indicated by lines of increasing thickness.

R:FR ratios have minimal impact on phenotype, while CBGa is upregulated

A modest increase in plant height was observed in the high R:FR group of plants, but due to the harvest of the terminal meristems for qPCR and HPLC analysis, accurate plant height could not be recorded. Measurement of internodal length, leaf surface area, and leaf, stem, and flower biomass showed no significant differences (Fig. 6).

However by week 6, cannabigerolic acid (CBGa) levels in the fresh flower were more than 20% greater in the high R:FR group than the low, while end-product cannabinoid levels remained unchanged (Fig. 7). This pattern continued through harvest and drying, where in the composite dry-weight analysis, we see an approximately 25% increase of CBGa in the high RFR treatment group, but no change in any of the other cannabinoids. This
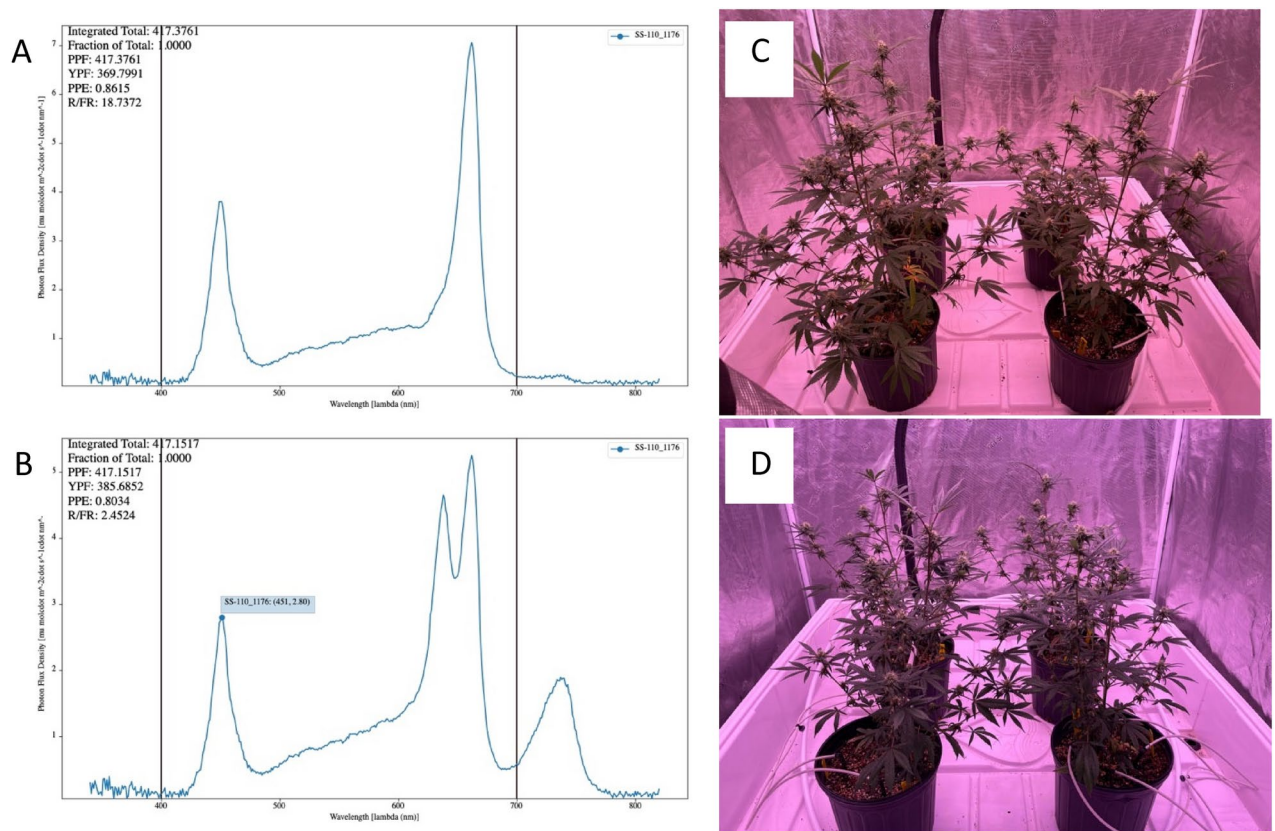
Fig. 4. Experimental Design and Validation of R:FR treatments. A and B Spectroradiographs confirming the R:FR ratios generated by the light fixtures. High R:FR conditions show a large peak at 660 nm and no peak in the 700 nm + (FR) range (Panel **A**). Low R:FR conditions show a large peak at 660 nm as well as a large peak in the FR 700 nm + range (Panel **B**).C and D: Plants in their treatment conditions, the light color is determined by treatment. The high R:FR light appears pinker (Panel **C**). The low R:FR light appears more purple (Panel **D**). In both conditions, PPFD and B:R ratios are normalized. Spectral conditions calculated from 5 measurement zones. High R:FR Conditions: R:FR Ratio = 19.27, standard deviation of 0.48. Average R:B Ratio = 1.91, standard deviation of 0.04. Average PFD = 399.34, standard deviation of 6.08. Low R:FR Condtiions (Control): R:FR Ratio = 2.65 standard deviation of 0.09, Average R:B Ratio = 1.88, standard deviation of 0.03, Average PFD = 398.41, standard deviation of 7.96.

indicates that the high RFR treatment increased the expression of CBGa, but that CBGa was not the rate-limiting reactant for end-product cannabinoid synthesis in this example.

In the graphical illustration below (Fig. 8) we propose a putative model for the activation of FT and floral meristem identity through a phytochrome-mediated response to SD photoperiod involving PIFs, WNK kinases, and pseudo-response regulators.
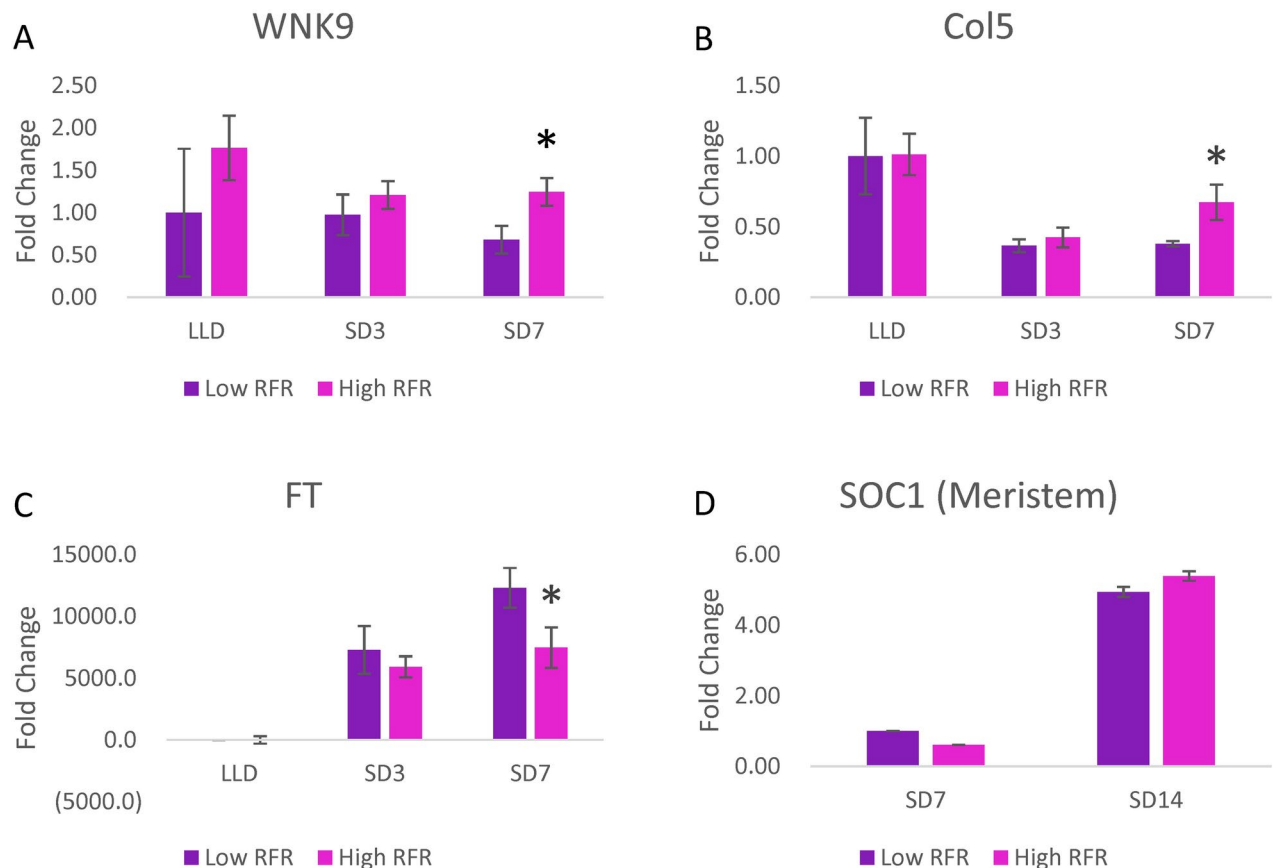
Methods
Plant growth
"Wife" cultivar hemp used for the experiment, cuttings sourced from internal stock. Plants were grown inside of Gorrilla Grow grow tent structures, each outfitted with batch-tank fertigation. Humidity and temperature were controlled via high-output ventilation fans and thermostats. Cuttings were stuck in Rockwool using Clonex Root Gel, allowed to root for two weeks under high pressure sodium lighting, and then planted in #900 blow molded nursery pots in ProMix BX-25 supplemented with 3 g/L Osmocote 15–9-12 (ICL Specialty Fertilizers). Cuttings were then placed into grow tents and allowed to grow vegetatively for another two weeks before switching to flower. Plants were fed via drip fertigation CLF Jack's 15–30-15 (JR Peters, Inc).

Design of spectral conditions
Adaptiiv ATG-1000 Liquid-Cooled LED light fixtures were used to create the spectral conditions for each treatment. Blue, Red, Far Red, and White illuminance was adjusted to achieve the target R:FR ratio while normalizing B:R ratios and PPFD. The resulting spectral composition was measured using an Apogee SS-110 Spectroradiometer, and a composite of 5 measurements was taken per treatment area to determine the variability.

Fig. 5. qPCR data describing the fold change in gene expression in cannabis leaves across SD transition under high and low R:FR conditions. 4 Biological replicates per assay. UB used as reference, delta-delta Ct method used to calculate relative expression change. In each assay, the last long day is set as control (1) except SOC1, in which case Day 7 is the control. Although no change was observed from the treatment by day 3, by the seventh short day, three genes in the study were affected by high RFR ratios: WNK9 (Panel **A**), COL5 (Panel **B**), and FT (Panel **C**). Expression of WNK9 and COL5 both increased significantly. FT expression was significantly reduced, but this downregulation had no impact on downstream SOC1 (Panel **D**). Student's t-test was utilized for calculating statistical significance of means separation. One star (*) indicates a P-value below 0.05, while two stars (**) indicates a P-value below 0.005.

Sampling

Leaf samples were taken immediately before switching to short days (LLD), and the follwing day was designated "Short Day 1" (SD1). The fifth, fourth, third, second and first fully expanded leaf from the terminal meristem was selected for each consecutive sampling, in that order. Leaf samples taken again at SD 1, 3, 7, and 14, at EOD appx 2 h before lights out. Transitioning floral meristem samples taken at days 7 and 14 from the 2nd and 1st most dominant meristem respectively. Flower samples for RNA taken at SD 28. Flower samples taken for FWT cannabinoids from remaining undamaged terminal phytomers at SD 28 and 42. Final cannabinoid sample taken from dry composite of each plant.

RNA-Seq library preparation and sequencing

RNA isolated using NucleoSpin Plant and Fungi RNA Isolation Kit. RNA quality and integrity were assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies) and RNA with an RNA integrity number (RIN) ≥ 7 was used for library preparation. RNA-Seq libraries were prepared using the NEBNext Ultra II RNA Library Prep Kit for Illumina, New England Biolabs according to the manufacturer's instructions. Briefly, mRNA was isolated using oligo(dT) magnetic beads, fragmented, and reverse-transcribed into cDNA. Double-stranded cDNA was end-repaired, A-tailed, and ligated to sequencing adapters, followed by PCR amplification. The libraries were quantified using Qubit dsDNA HS Assay Kit, Thermo Fisher Scientific and their size distributions were assessed with the Agilent Bioanalyzer.Sequencing was performed on an Illumina HiSeq 4000 platform in 100 bp paired-end mode on a single lane, achieving an average read depth of approximately 20 million reads per sample. Four biological replicates were analyzed for each condition, yielding a total of 8 samples and a total of > 160 million reads reads across all samples.
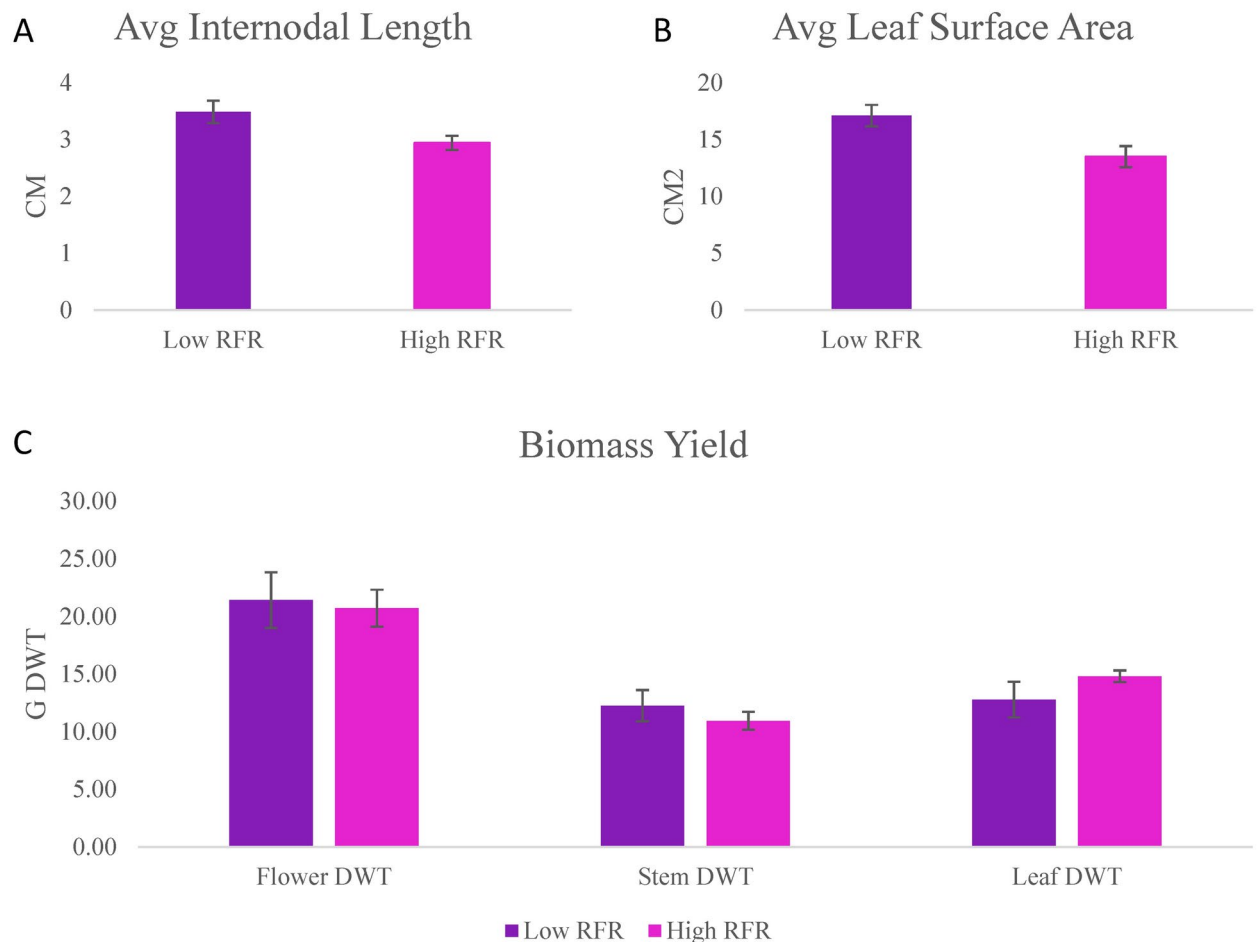
Fig. 6. Phenotype impact measurements. Measurements were taken on day 42 immediately prior to harvest. No change in average internodal length was observed (Panel **A**). Leaf surface area was also unaffected (Panel **B**). Whole plants were hung to dry for 14 days at 40% humidity and 60F until they had reached a moisture content of ~ 10%. Tissue types were separated and weighed (Panel **C**). No significant differences observed.

RNA-seq analysis

Raw sequencing data were assessed for quality using *FastQC* (version 0.12.0) and trimmed to remove adapter sequences and low-quality bases using *Trimmomatic* (version 0.39) with the standard parameters[44]. High-quality reads were aligned to the reference transcripts for 9 different Cannabis reference transcriptomes. Since Cannabis is still in its infancy in the information available for genomic information and short read RNA quantification is directly correlated to the quality of its reference, assessment was given based on mapping rate. From NCBI, the transcriptomes for PRJEB21674, PRJNA73819, PRJNA245084, PRJNA435671, PRJNA483805, PRJNA498707 (contained 4 genomes). Salmon (v1.10.3) was used for alignment and indexes were developed using these reference transcriptomes[45]. Transcript level quantification (TPM) was used then imported into *DESeq2* (version 1.44.0) using tximport and DESeqDataSetFromTximport[46]. LFC shrinkage[47] was used as it improved the noise and preserved large differences (supplemental Fig. 3 and 4). Differential gene expression analysis was carried out with *DESeq2*, applying a false discovery rate (FDR) of < 0.05 to identify significantly differentially expressed genes.

Protein–protein interaction prediction

STRING version 12.0 used for PPI predictions. Physical subnetwork type, sourced from textmining, experiments, databases, co-expression, neighborhood, gene fusion and co-occurrence. Line thickness indicates strength of data support. Minimum required interaction score set to medium confidence (0.400). Jamaican Lion proteome used to generate reference STRG0A70MMF. Input amino acid sequences derived from nucleotide sequences discovered in DESeq.

HPLC–UV

For each analysis, ~ 100 mg of flower tissue was snap frozen in liquid nitrogen and milled using a Tissuelyzer swing-mill. Each sample was then extracted into 10 ml methanol measured with a volumetric flask, filtered using
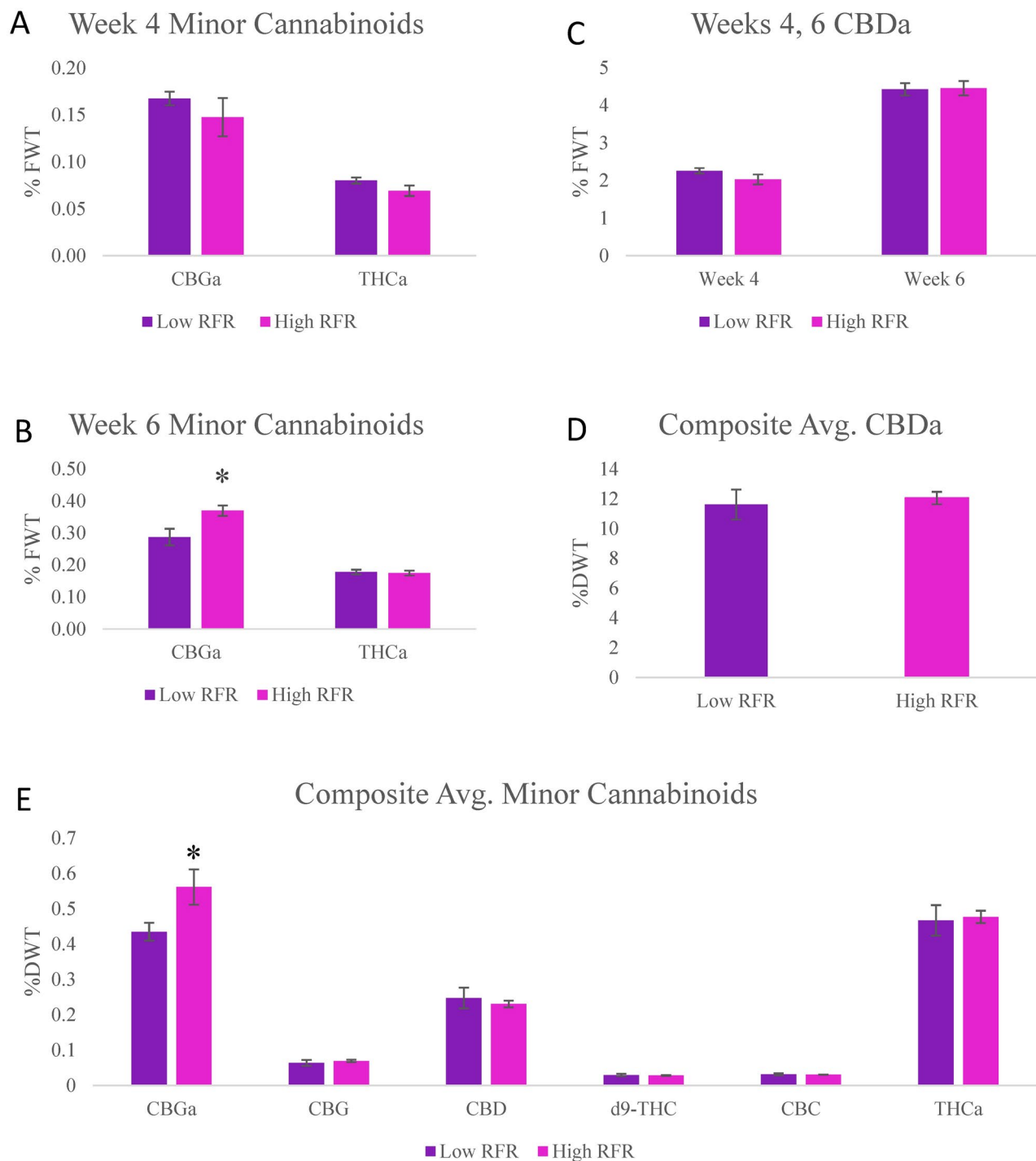
Fig. 7. HPLC–UV results for cannabinoid content in fresh and dry flower samples. Panel A depicts minor cannabinoids extracted from fresh flowers during the fourth week of flowering. No change is observed from the high R:FR treatment. Panel B depicts minor cannabinoids extracted from fresh flowers during the sixth week of flowering. CBGa has been significantly increased in the high R:FR group. Panel C demonstrates that CBDa remained unchanged at weeks 4 and 6. Panel D depicts avg. CBDa extracted from the dried and milled composite of flowers from each individual replicate: no change is observed. Panel E depicts the minor cannabinoids extracted from these same samples. The increased CBGa persisted through week 6, and into the dried flower. Student's t-test was utilized for calculating statistical significance of means separation. One star (*) indicates a P-value below 0.05, while two stars (**) indicates a P-value below 0.005.
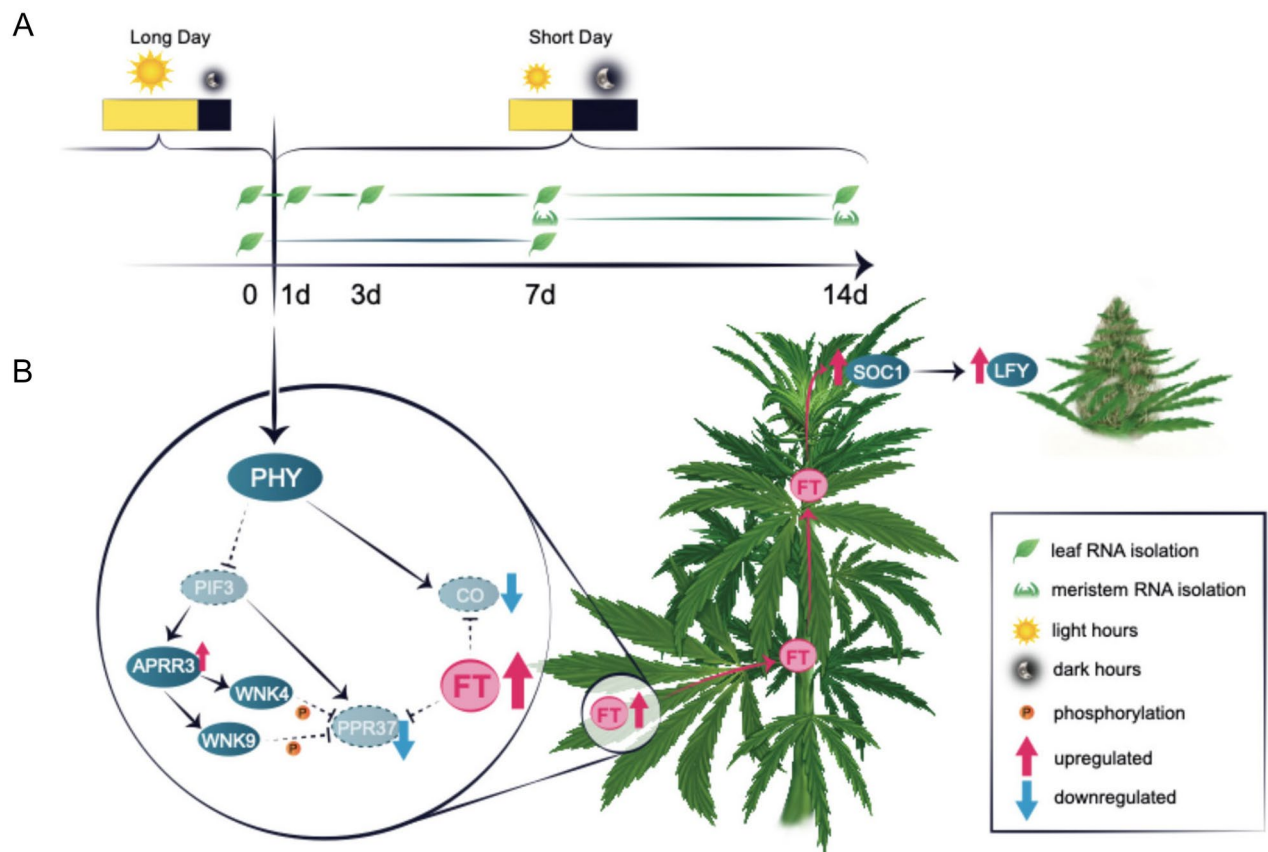
Fig. 8. A putative model for the transcriptional activation of floral meristem identity by PHY via FT in cannabis, an SD plant. The model summarizes the data presented here, suggesting that PHY integrates photoperiod entrainment, targeting PIF for degradation, which stimulates a pseudo-response regulator/WNK kinase signaling cascade. Ultimately, WNK kinases phosphorylate PRR37, and its inhibition allows for the expression of FT, a mobile signal which is translocated to the developing vegetative meristem. FT activates its downstream target SOC1 in the meristem, where floral initiation genes are activated. **A**: Experimental Design. **B**: Putative model of interactions. Dotted lines indicate degradation while solid lines/ arrows indicate functional relationship. Figure illustration produced by Daria Chrobok, DC SciArt.

a 0.45 µm nylon syringe filter, diluted $10\times$ using a Hamilton glass syringe and analyzed using Shimadzu Hemp Analyzer High Sensitivity Package according to manufacturer's protocol[48]. Calibration curve was produced immediately prior to analysis. Autosampler loaded with one standard QC vial for every 6 sample vials, as well as a single methanol blank.

Internodal length, leaf SA and biomass
The 3rd branch from the terminal meristem was removed and the 1st, 2nd and 3rd leaves/ internodes were measured. Leaf SA measured using LeafScan app. Flowers dried using commercial methods and water weight calculated.

qPCR
Ubiquitin used as housekeeping gene for all reactions[49]. Primers designed using NCBI's primer design tool. RNA isolated using NucleoSpin Plant and Fungi RNA Isolation Kit. cDNA synthesized using iScript Reverse transcriptase. Bio-Rad CFX-Connect Real time qPCR system was used with iTaq SYBR Green Universal Master Mix. Relative expression values calculated using delta-delta Ct method. Student's t-tests were used to establish statistical significance. One-way ANOVA and Tukey HSD test performed in Rstudio.

Data availability
All raw and processed sequencing data are available through SRA under accession number SUB14928528. Custom analysis scripts used in this study are available on Github at https://github.com/bioinfonerd/cdl_cannabis.

References

1. Cheng, M.-C., Kathare, P. K., Paik, I. & Huq, E. Phytochrome Signaling Networks. *Annu. Rev. Plant Biol.* **72**, 217–244 (2021).
2. Igamberdiev, A. U., Eprintsev, A. T., Fedorin, D. N. & Popov, V. N. Phytochrome-mediated regulation of plant respiration and photorespiration. *Plant, Cell Environ.* **37**, 290–299 (2014).
3. Li, J., Li, G., Wang, H. & Wang Deng, X. Phytochrome Signaling Mechanisms. *The Arabidopsis. Book* **9**, e0148 (2011).
4. Seaton, D. D. et al. Dawn and photoperiod sensing by phytochrome A. *Proc Natl Acad Sci USA* **115**, 10523–10528 (2018).
5. Franklin, K. A. & Whitelam, G. C. Phytochromes and shade-avoidance responses in plants. *Ann. Bot.* **96**, 169–175 (2005).
6. Ashdown, I. Phytochrome and PSS: 'Think Beyond Pink'. (2016).
7. Sager, J. C., W. O. Smith, J. L. Edwards, and K. L. Photosynthetic Efficiency and Phytochrome Equilibria Determination Using Spectral Data. *Transactions of the ASAE* **31**, (1988).
8. Jaeger, K. E. & Wigge, P. A. FT protein acts as a long-range signal in arabidopsis. *Curr. Biol.* **17**, 1050–1054 (2007).
9. Hall, J., Bhattarai, S. P. & Midmore, D. J. Review of flowering control in industrial hemp. *J. Nat. Fibers* **9**, 23–36 (2012).
10. Lisson, S. N., Mendham, N. J. & Carberry, P. S. Development of a hemp (Cannabis sativa L.) simulation model 2The flowering response of two hemp cultivars to photoperiod. *Aust. J. Exp. Agric.* **40**, 413 (2000).
11. Ishikawa, R., Shinomura, T., Takano, M. & Shimamoto, K. Phytochrome dependent quantitative control of Hd3a transcription is the basis of the night break effect in rice flowering. *Genes Genet. Syst.* **84**, 179–184 (2009).
12. Kurtz, L. E., Brand, M. H. & Lubell-Brand, J. D. Gene dosage at the autoflowering locus effects flowering timing and plant height in triploid cannabis. *J. Amer. Soc. Hort. Sci.* **148**, 83–88 (2023).
13. Teotia, S. & Tang, G. To Bloom or not to bloom: Role of MICRORNAS in plant flowering. *Mol. Plant* **8**, 359–377 (2015).
14. Shim, J. S., Kubota, A. & Imaizumi, T. Circadian clock and photoperiodic flowering in arabidopsis: CONSTANS is a hub for signal integration. *Plant Physiol.* **173**, 5–15 (2017).
15. Mizoguchi, T. et al. Distinct roles of *GIGANTEA* in promoting flowering and regulating circadian rhythms in arabidopsis. *Plant Cell* **17**, 2255–2270 (2005).
16. Duek, P. D. & Fankhauser, C. HFR1, a putative bHLH transcription factor, mediates both phytochrome A and cryptochrome signalling. *Plant J.* **34**, 827–836 (2003).
17. King, R. W., Hisamatsu, T., Goldschmidt, E. E. & Blundell, C. The nature of floral signals in Arabidopsis I Photosynthesis and a far-red photoresponse independently regulate flowering by increasing expression of FLOWERING LOCUS T (FT). *J. Exper. Botany* **59**, 3811–3820 (2008).
18. Wenkel, S. et al. CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of *Arabidopsis. Plant Cell* **18**, 2971–2984 (2006).
19. Leckie, K. M. *et al. Loss of Daylength Sensitivity by Splice Site Mutation in Cannabis.* http://biorxiv.org/lookup/doi/https://doi.org/10.1101/2023.03.10.532103 (2023) https://doi.org/10.1101/2023.03.10.532103.
20. Murakami-Kojima, M., Nakamichi, N., Yamashino, T. & Mizuno, T. The APRR3 component of the clock-associated APRR1/TOC1 quintet is phosphorylated by a novel protein kinase belonging to the WNK family, the gene for which is also transcribed rhythmically in Arabidopsis thaliana. *Plant Cell Physiol.* **43**, 675–683 (2002).
21. Sun, C. et al. Bifunctional regulators of photoperiodic flowering in short day plant rice. *Front. Plant Sci.* **13**, 1044790 (2022).
22. Yano, M. et al. Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the arabidopsis flowering time gene CONSTANS. *Plant Cell* **12**, 2473–2483 (2000).
23. Copley, T. R., Duceppe, M.-O. & O'Donoughue, L. S. Identification of novel loci associated with maturity and yield traits in early maturity soybean plant introduction lines. *BMC Genomics* **19**, 167 (2018).
24. Watanabe, S. et al. A map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering. *Genetics* **188**, 395–407 (2011).
25. Pan, G. et al. Genome-wide identification, expression, and sequence analysis of CONSTANS-like gene family in cannabis reveals a potential role in plant flowering time regulation. *BMC Plant Biol* **21**, 142 (2021).
26. Dowling, C. A. *et al. A* FLOWERING LOCUS T *Ortholog Is Associated with Photoperiod-Insensitive Flowering in Hemp (* Cannabis Sativa *L.).* http://biorxiv.org/lookup/doi/https://doi.org/10.1101/2023.04.21.537862 (2023) https://doi.org/10.1101/2023.04.21.537862.
27. Steel, L., Welling, M., Ristevski, N., Johnson, K. & Gendall, A. Comparative genomics of flowering behavior in Cannabis sativa. *Front. Plant Sci.* **14**, 1227898 (2023).
28. Spitzer-Rimon, B. et al. Non-photoperiodic transition of female cannabis seedlings from juvenile to adult reproductive stage. *Plant Reprod* **35**, 265–277 (2022).
29. Chen, X. et al. Whole-genome resequencing of wild and cultivated cannabis reveals the genetic structure and adaptive selection of important traits. *BMC Plant Biol* **22**, 371 (2022).
30. Spitzer-Rimon, B., Duchin, S., Bernstein, N. & Kamenetsky, R. Architecture and florogenesis in female cannabis sativa plants. *Front. Plant Sci.* **10**, 350 (2019).
31. Das, A., Chaudhury, S., Kalita, M. C. & Mondal, T. K. In silico identification, characterization and expression analysis of miRNAs in Cannabis sativa L. *Plant Gene* **2**, 17–24 (2015).
32. Wu, G. et al. The sequential action of miR156 and miR172 regulates developmental timing in arabidopsis. *Cell* **138**, 750–759 (2009).
33. Aukerman, M. J. & Sakai, H. Regulation of flowering time and floral organ identity by a MicroRNA and Its *APETALA2* -like target genes. *Plant Cell* **15**, 2730–2741 (2003).
34. Xie, K., Wu, C. & Xiong, L. Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and microRNA156 in rice. *Plant Physiol.* **142**, 280–293 (2006).
35. Chen, X. A MicroRNA as a translational repressor of *APETALA2* in *Arabidopsis* flower development. *Science* **303**, 2022–2025 (2004).
36. Park, W., Li, J., Song, R., Messing, J. & Chen, X. CARPEL FACTORY, a Dicer Homolog, and HEN1, a novel protein, act in microRNA metabolism in arabidopsis thaliana. *Curr. Biol.* **12**, 1484–1495 (2002).
37. Hu, J. et al. Unravelling miRNA regulation in yield of rice (Oryza sativa) based on differential network model. *Sci. Rep.* **8**, 8498 (2018).
38. McKernan, K. *et al.* Cryptocurrencies and Zero Mode Wave guides: An unclouded path to a more contiguous Cannabis sativa L. genome assembly. Preprint at https://doi.org/10.31219/osf.io/7d968 (2018).
39. Zhang, B., Wang, L., Zeng, L., Zhang, C. & Ma, H. *Arabidopsis* TOE proteins convey a photoperiodic signal to antagonize CONSTANS and regulate flowering time. *Genes Dev.* **29**, 975–987 (2015).
40. Du, A. et al. The DTH8-Hd1 module mediates day-length-dependent regulation of rice flowering. *Mol. Plant* **10**, 948–961 (2017).
41. Toth, J. A., Stack, G. M., Carlson, C. H. & Smart, L. B. Identification and mapping of major-effect flowering time loci Autoflower1 and Early1 in Cannabis sativa L. *Front. Plant Sci.* **13**, 991680 (2022).
42. Xia, Z. et al. Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering. *Proc. Natl. Acad. Sci. U.S.A.* https://doi.org/10.1073/pnas.1117982109 (2012).
43. Pumplin, N. et al. DNA methylation influences the expression of *DICER-LIKE4* Isoforms, which encode proteins of alternative localization and function. *Plant Cell* **28**, 2786–2804 (2016).
44. Williams, C. R., Baccarella, A., Parrish, J. Z. & Kim, C. C. Trimming of sequence reads alters RNA-Seq gene expression estimates. *BMC Bioinformatics* **17**, 103 (2016).

45. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417–419 (2017).
46. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
47. Zhu, A., Ibrahim, J. G. & Love, M. I. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics* **35**, 2084–2092 (2019).
48. Shimadzu Corporation. The Potency Determination of 15 Cannabinoids using the Cannabis Analyzer for Potency. (2020).
49. Deguchi, M. et al. Selection and validation of reference genes for normalization of qRT-PCR data to study the cannabinoid pathway genes in industrial hemp. *PLoS ONE* **16**, e0260660 (2021).

Acknowledgements

Author contributions

S.H. Designed and executed all experiments except for transcriptome mapping and DESeq, which was performed by N.J., and RNA sequencing, which was performed by UConn's Center for Genome Innovation. All figures and tables prepared by S.H. except Supplemental Figs. 1, 2, 3, and 4, which were prepared by N.J. Manuscript was written and edited by S.H. G.B. and N.J. reviewed the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-00430-7.

Correspondence and requests for materials should be addressed to G.A.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.