



OPEN ACCESS

EDITED BY

Dilip R. Panthee,
North Carolina State University,
United States

REVIEWED BY

Abhinandan Surgonda Patil,
International Rice Research Institute
(IRRI), Philippines
Alicja Macko-Podgórn,
University of Agriculture in Kraków,
Poland

*CORRESPONDENCE

Lawrence B. Smart
lbs33@cornell.edu

†PRESENT ADDRESS

Craig H. Carlson,
Cereal Crops Research Unit, Edward T.
Schafer Agricultural Research Center,
USDA, ARS, Fargo, ND, United States

SPECIALTY SECTION

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

RECEIVED 11 July 2022

ACCEPTED 22 August 2022

PUBLISHED 21 September 2022

CITATION

Toth JA, Stack GM, Carlson CH and
Smart LB (2022) Identification and
mapping of major-effect flowering
time loci *Autoflower1* and *Early1* in
Cannabis sativa L.
Front. Plant Sci. 13:991680.
doi: 10.3389/fpls.2022.991680

COPYRIGHT

© 2022 Toth, Stack, Carlson and
Smart. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Identification and mapping of major-effect flowering time loci *Autoflower1* and *Early1* in *Cannabis sativa* L.

Jacob A. Toth, George M. Stack, Craig H. Carlson[†] and
Lawrence B. Smart^{*}

Horticulture Section, School of Integrative Plant Science, Cornell University, Cornell AgriTech,
Geneva, NY, United States

Flowering time is an important trait for all major market classes of hemp (*Cannabis sativa*), affecting yields and quality of grain, fiber, and cannabinoids. *C. sativa* is usually considered a short-day plant, flowering once night length reaches a critical threshold. Variations in flowering time within and across cultivars in outdoor grown populations have been previously identified, likely corresponding to genetic differences in this critical night length. Further, some *C. sativa* are photoperiod insensitive, colloquially referred to as “autoflowering.” This trait has anecdotally been described as a simple recessive trait with major impacts on phenology and yield. In this work, the locus responsible for the “autoflower” trait (*Autoflower1*), as well as a major-effect flowering time locus, *Early1*, were mapped using bulked segregant analysis. Breeder-friendly high-throughput molecular marker assays were subsequently developed for both loci. Also detailed are the flowering responses of diverse cultivars grown in continuous light and the result of crossing two photoperiod insensitive cultivars of differing pedigree.

KEYWORDS

hemp (*Cannabis sativa* L.), bulk segregant analysis (BSA), flowering time, autoflower, early flowering, continuous light

Introduction

Hemp (*Cannabis sativa*) is a multi-use crop that is widely considered a photoperiod sensitive, short-day plant. Flowering time is important for all major market classes, and uniform flowering dates within a cultivar are essential for ease of harvest. Fiber hemp benefits from a long growing season, as harvest usually occurs around the flowering date, and early flowering results in a shorter vegetative growth phase to accumulate biomass (Salentijn et al., 2019). Grain hemp must flower early enough such that grain can be harvested before frost if growing in temperate latitudes, but precocious flowering can

lead to severe yield penalties due to a lack of time to accumulate biomass that provides photosynthate for grain filling. This is especially challenging for subtropical and tropical latitudes where most days of the year have nights longer than the critical night threshold (Zhang et al., 2021). For cannabinoid production, as with grain production, precocious flowering may result in reduced floral biomass yield, while plants that do not flower by the end of the season fail to accumulate high concentrations of cannabinoids (Stack et al., 2021). Additionally, cannabinoid profiles change throughout the maturation of the inflorescence, making initiation of flowering an important factor in timing regulatory compliance testing and harvest (Toth et al., 2021; Zhang et al., 2021).

Previous work has outlined differences in photoperiod threshold across different cultivars, in controlled environments as well as field conditions (Stack et al., 2021; Zhang et al., 2021). It has also been well established that some plants are photoperiod insensitive (day neutral), a trait proposed to have been introgressed from high-latitude populations, which have been classified by some as a putative species, *Cannabis ruderalis* (McPartland, 2018). Photoperiod insensitivity is sometimes referred to colloquially as “autoflower” (Gloss, 2015). This trait has been suggested to be inherited in a simple, recessive, Mendelian fashion, but there are limited data on this in the peer-reviewed literature (Green, 2005). A patent covering molecular markers and biotechnological manipulation of genes responsible for “autoflower” is held by Phylos Biosciences (Phylos Bioscience, International Patent WO 2021/097496 A2). In addition to commercial high-cannabinoid “autoflower” cultivars, several grain cultivars, such as ‘FINOLA,’ have been referred to as “autoflowering” in the literature (Van Bakel et al., 2011), but have a distinct phenotype compared to photoperiod-insensitive high-cannabinoid cultivars in that the height of mature ‘FINOLA’ depends greatly on latitude (being shorter at lower latitudes), while “autoflower” high-cannabinoid cultivars do not appear to exhibit this phenomenon (Callaway, 2002; Yang et al., 2020; Stack et al., 2021). However, it is not clear if the genetic mechanism for photoperiod insensitivity is the same in all cultivars (Zhang et al., 2021). Diagnostic molecular markers and complementation assays could help resolve this question.

It is a well-established phenomenon that there is significant population structure in *C. sativa*, associated at least in part with recent breeding history and geography (Carlson et al., 2021). While there is ongoing debate on the specifics of the nature of this population structure (hindered in part because of the ease of intercrossing between subgroups of *C. sativa*), there is strong support for at least two subpopulations, which have been described as subspecies (McPartland, 2018). The two subspecies that have been described differ in end use and likely origin, with *C. sativa* ssp. *sativa* grown for grain and fiber originally in northern European latitudes and *C. sativa* ssp. *indica* grown for cannabinoid production

originally in Southeast Asia, including India. Various other subpopulations have been described, including a distinct clade of *C. sativa* with geographic origins in China (Carlson et al., 2021; Ren et al., 2021). Different taxonomic classifications have also been proposed, including the putative species *C. ruderalis*, which has been considered the source of the “autoflower” trait in all *C. sativa* populations (Green, 2005).

Genetic pathways for the induction of flowering are largely conserved across dicot plants, with major photosensory, thermosensory, and age-related pathways converging on major floral integrator genes (Mouradov et al., 2002; Jung and Müller, 2009). These floral integrator genes, including CO, FT, SOC1, and FLC, result in the expression of floral meristem identity genes such as AP1 and LFY (Mouradov et al., 2002). The expression of these floral meristem identity genes result in a switch from a vegetative phase to a reproductive phase. Upstream of the major floral integrator genes are a host of other well-established genes including PRR37, casein kinase I, AP2 group genes, and others, which have been associated with flowering time in various plants including rice (Hori et al., 2013), sorghum (Murphy et al., 2011), and pepper (Yuan et al., 2021).

There has been extensive research into genes involved in photoperiod response and earliness *per se* in a variety of crops, aided by extensive work in Arabidopsis (Blümel et al., 2015). However, there has been relatively little research in this area in *C. sativa*. A recent genome-wide association study (GWAS) in fiber hemp implicated genes in the major photosensory, thermosensory, and age-related pathways as well as a host of transcription factors in time to flower (Petit et al., 2020). While the genes implicated in this study serve as a potential starting point, more study on the genetics of flowering time control in *C. sativa* is required for predictive breeding efforts. The conserved nature of flowering induction may ease the discovery of relevant genes in extant germplasm.

Our group previously identified several populations marketed as F₁ hybrids (‘Umpqua’ and ‘Deschutes’) that have individuals with two distinct flowering times, approximately a month apart (Stack et al., 2021). This segregation ratio would be expected if one parent was heterozygous for a major effect flowering time locus while the other parent was homozygous at that locus. If this were the case, such a significant locus would be well suited for development of high-throughput molecular markers such as PACE (PCR Allele Competitive Extension, 3CR). As differences in flowering are not obvious when the plant is in an early vegetative state, early screening with molecular markers for this trait or the autoflower trait could be very useful (Toth et al., 2018).

For essentially qualitative traits controlled by major effect flowering time loci as described here, bulk segregant analysis (BSA) has been successfully used to map genes and generate molecular markers and related assays (Song et al., 2017). Bulk segregant analysis is a technique that utilizes the sequencing of

pooled DNA samples from individuals with the same phenotype in contrasting groups in a segregating population, and has been used effectively in a range of crops including *C. sativa* (Ban and Xu, 2020; Welling et al., 2020). Bulk segregant analysis usually involves short-read sequencing and subsequent alignment to a reference genome, but BSA involving long reads and reference-free techniques have been developed (Nordström et al., 2013; Segawa et al., 2021). The number of individuals in the pools must be sufficiently large to randomize the association of all regions of the genome except the region or regions associated with the trait of interest. Compared with other methods of mapping, this technique has the advantage of obtaining whole genome sequences of the region of interest, alleviating the issue of ascertainment bias present in other methods of sequencing mapping populations such as single nucleotide polymorphism (SNP) chips or genotyping-by-sequencing (GBS). It is also cheaper and results in higher read depth than individually sequencing genomes, but multiple sequencing efforts would be required for mapping more than one trait. Bulk segregant analysis can also be conducted using pre-defined molecular markers instead of direct sequencing, but the decreasing cost of sequencing has made these approaches less common (Zhang et al., 2009; Becker et al., 2011).

Once sequencing data are obtained from contrasting pools, a comparison of regions that differ in allelic frequency can be performed (Magwene et al., 2011). In the case of a simple recessive trait in an F_2 population, one pool should be homozygous for a region containing the causative gene, while the other pool should have an alternative allelic frequency of ~33% in that region. In the case of a major gene in a backcross, one pool should be homozygous in a region and the other pool should be heterozygous. The difference between allele frequencies can be represented in a number of ways, including comparing the number of significantly different SNPs in a region determined through Fisher's Exact Test, a G-Test statistic, or through the delta-allele frequency method (also called the delta-SNP method) which involves determining the difference in allele frequency directly (Zhang and Panthee, 2020).

Materials and methods

Field and greenhouse trials of populations segregating for flowering time

A population segregating for photoperiod insensitivity was developed by first crossing a female autoflower plant from a feminized (all-female) seed lot numbered KG9202 (generously provided by Kayagene, Hollister, CA, United States) with a late flowering, photoperiod sensitive 'Otto II' plant (generously provided by Edgar Winters, WinterFox Farms, Klamath Falls, OR) determined to be male and cannabinoid chemotype III

using molecular markers (Toth et al., 2020) to produce F_1 family GVA-H-19-1148. These parental cultivar populations were previously trialed in the 2019 Cornell high-cannabinoid hemp field trial (Stack et al., 2021). One selected photoperiod-sensitive female F_1 plant (GVA-H-19-1148-002) was multiplied by rooting stem cuttings, then one ramet was treated with silver thiosulfate to induce male flowers that pollinated multiple genetically identical female plants to generate F_2 seed labeled GVA-H-20-1080 (Carlson et al., 2021). This technique results in an entirely female, or feminized, population (Lubell and Brand, 2018).

A second population segregating for photoperiod insensitivity was 'TJ's CBG' (generously provided by Stem Holdings Agri, Eugene, OR), which was evaluated in the 2020 Cornell CBG hemp field trials and displayed a CBG-dominant chemotype (chemotype IV).

For initial assessment of photoperiod insensitivity in the segregating populations, seeds of each were sown in potting mix in 50-cell SureRoot trays on December 16, 2019 and grown in a greenhouse with a 16L:8D light schedule. Eighty-eight healthy plants in each population were transplanted to one-gallon pots on February 3, 2020. While flowering was evident on some plants at this point, rating for terminal flowering as previously defined (Stack et al., 2021) was completed on March 23, 2020.

The high-CBD cultivar 'Umpqua' (generously provided originally by Industrial Seed Innovations) was evaluated in the Cornell high-cannabinoid hemp field trials in 2019 and 2020 where flowering time was carefully assessed. Details about the 2019 trial are available in Stack et al. (2021). The 2020 trial was executed using similar protocols, but with a different seed lot of 'Umpqua' generously provided by Arcadia Biosciences (Davis, CA). An additional 100 plants taken equally from both seed lots were transplanted outdoors on July 22, 2021 in a trial to evaluate flowering time in Geneva, NY using similar protocols. The 2021 flowering time field trial also included 96 individuals from the KG9202 × 'Otto II' F_2 population GVA-H-20-1080 and 26 plants of 'Hempress' (generously provided by Point3 Farma, Center, CO). Height and wet biomass was recorded for each plant in population GVA-H-20-1080 in this trial. Additional populations segregating for photoperiod insensitivity were identified in the 2020 Cornell CBG hemp field trial.¹

To evaluate photoperiod insensitivity across diverse germplasm, 50 seeds of each population were sown in potting mix in 50-cell SureRoot trays on April 20, 2021 unless otherwise noted and grown in a greenhouse under continuous supplemental lighting from high pressure sodium lamps. Flowering was assessed weekly. Male flowering was considered to have started when the length of internodes at the apex

¹ <https://hemp.cals.cornell.edu/resource/2020-cbg-hemp-cultivar-trial/>

of the plant shortened and male buds were clearly visible at the growing tip.

A complementation cross was completed between two photoperiod-insensitive plants: pollen parent ‘Picolo’ (generously provided by Hemp Genetics International, Saskatoon, SK), and a homozygous *Autoflower1/Autoflower1* seed parent from the ‘Le Crème’ cultivar population (generously provided by Ventura Seed Company, Camarillo, CA, segregating 1:3 for photoperiod insensitivity). The F₁ plants from this cross were grown under 16L:8D with a 1 h night break in 50-cell SureRoot trays alongside known photoperiod sensitive and insensitive cultivars. Ten plants from each population were established on January 5, 2022. ‘Auto CBD’ and ‘Le Crème’ were feminized populations with no males. ‘RN16’ was a dioecious photoperiod-sensitive high CBD hemp cultivar (Stack et al., 2021).

Bulk segregant analysis sequencing

DNA was extracted using a Qiagen DNeasy 96-well kit from young leaf tissue collected from plants in population GVA-H-20-1080 and dried on silica gel. Two pools were created by combining equal amounts of DNA from 28 flowering, photoperiod-insensitive plants and 25 non-flowering, photoperiod-sensitive plants. Illumina TruSeq libraries with an insert size of ~500 bp were constructed for each pool by the Cornell Institute of Biotechnology then paired end 151 bp sequencing was performed on the Illumina NextSeq 2000 platform with ~35× coverage.

DNA was extracted from dried, milled floral biomass samples of ‘Umpqua’ as previously described (Toth et al., 2020) for 15 early flowering and 15 late-flowering plants from the 2019 and 2020 trials and from 15 early flowering and 15 late-flowering samples from the 2021 flowering time trial. Illumina TruSeq libraries were constructed for each phenological pool and sequenced on the Illumina NextSeq 2000 platform, as described above.

Reads were aligned to the CBDRx-cs10 (GCF_900626175.2) genome assembly (Grassa et al., 2021) using Geneious Prime software (Biomatters, Inc., San Diego, CA, United States) using the Geneious mapper algorithm at the fastest speed with three iterations. Variants were also called in the Geneious Prime environment (coverage > 3, minimum variant frequency > 0.05). Variant calls were exported and modified using a custom Python script to be compatible with PyBSASeq (Zhang and Panthee, 2020). PyBSASeq was run using the “BulksOnly” protocol, assuming an F₂ population structure for GVA-H-20-1080 and a backcross population structure for ‘Umpqua’ bulks. Significant SNPs and regions were calculated using Fisher’s Exact Test, a G-Test statistic, and the delta-allele frequency method.

PCR allele competitive extension (PACE) genotyping assays

PCR allele competitive extension assays were designed manually in the Geneious Prime environment. PACE reactions were run according to the product manual (3CR Bioscience Ltd., Essex, United Kingdom). Polymorphic SNP in the *Autoflower1* region identified as perfectly associated with photoperiod phenotype in GVA-H-20-1080 pools were converted to PACE markers (Table 1 and Supplementary Table 1). These were assayed across multiple populations including the individual plants that formed the pool, a field grown population of GVA-H-20-1080, cultivars segregating for flowering time when grown under field conditions, and diverse photoperiod-sensitive and photoperiod insensitive cultivars. Unless otherwise noted, 8 plants from each cultivar or population were tested. For wide-germplasm testing, the primer sets AUTO-2 and EARLY-1 were used.

Results

Autoflower1 photoperiod insensitivity is a recessive Mendelian trait

Two populations segregating for photoperiod insensitivity (GVA-H-20-1080 and ‘TJ’s CBG’) were planted under non-inductive, long day conditions (16L:8D dark). In the GVA-H-20-1080 population, 28/88 plants flowered (31.8%), and in the ‘TJ’s CBG’ population, 24/88 plants flowered (27.3%). These data are not significantly different from 25% of the plants flowering ($\chi^2 P > 0.05$), consistent with a recessive allele at a single locus we are designating *Autoflower1* that was homozygous in KG9202, heterozygous in the photoperiod sensitive F₁ progeny of KG9202 × ‘Otto II,’ and segregating 1:2:1 in the GVA-H-20-1080 F₂ and ‘TJ’s CBG’ populations. The parents of the GVA-H-20-1080 population were also grown under these conditions, with KG9202 flowering alongside the *Autoflower1/Autoflower1* homozygotes and ‘Otto II’ not being induced to flower.

Mapping of the *Autoflower1* locus

Bulk segregant analysis of Illumina sequence pools of photoperiod-sensitive and -insensitive individuals showed clear statistical significance for the G-Test statistic in a region of Chromosome 1 (NC_044371.1) for population GVA-H-20-1080 (Figure 1). 230420 SNPs were included on Chromosome 1 in this analysis. No other chromosome reached significance by this metric (Figure 1). The significant region associated with *Autoflower1* spanned 17.74–22.94 Mb on Chromosome 1, with highly significant genomic windows (G-Test statistic > 11.5) at 18.59–19.70 Mb.

TABLE 1 PACE primers designed for the *Autoflower1* (AUTO) and *Early1* (EARLY) loci.

Primer	Group/location	Primer sequence
AUTO-1	18464905	
FAM	WT	GAAGGTGACCAAGTTCATGCTATCCAGGGTCTGGCTTTAAAAA
HEX	AUTO	GAAGGTCGGAGTCAACGGATTATCCAGGGTCTGGCTTTAAAAA
REV		CCATAAAATGATAAGTACACTCTAC
AUTO-2	19701425	
FAM	WT	GAAGGTGACCAAGTTCATGCTTTGGACTTCACCAAATGAGCCC
HEX	AUTO	GAAGGTCGGAGTCAACGGATTTTGGACTTCACCAAATGAGCCT
REV		CTTCTAACCCCTTTCATGAATG
AUTO-3	19731625	
FAM	AUTO	GAAGGTGACCAAGTTCATGCTCACAAGAATAATGCCCAAGAT
HEX	WT	GAAGGTCGGAGTCAACGGATTTCACAAGAATAATGCCCAAGAC
REV		CCTAGGTTGACATAGCCACCA
AUTO-4	19991224	
FAM	AUTO	GAAGGTGACCAAGTTCATGCTTCTCACTTCTGTCTTTTCCCT
HEX	WT	GAAGGTCGGAGTCAACGGATTCTCACTTCTGTCTTTTCCCC
REV		TCACAGTCTCAACAGGAGTGG
AUTO-5	21536161	
FAM	AUTO	GAAGGTGACCAAGTTCATGCTTTTTCATTTTCGGTGGGGTTTC
HEX	WT	GAAGGTCGGAGTCAACGGATTTTTTCATTTTCGGTGGGGTTT
REV		GGTTGGATGTTTCAGCTGAAG
EARLY-1	41445929	
FAM	EARLY	GAAGGTGACCAAGTTCATGCTGGATACTAGCCACTAGAAAGGTTT
HEX	WT	GAAGGTCGGAGTCAACGGATTGGATACTAGCCACTAGAAAGGTTG
REV		CGAAGGAGATAAAGACTGTGAG
EARLY-2	46288769	
FAM	EARLY	GAAGGTGACCAAGTTCATGCTATGTGTGTGTGCCTGTAGAACC
HEX	WT	GAAGGTCGGAGTCAACGGATTATGTGTGTGTGCCTGTAGAACC
REV		GTCCTAACCTTCAGAAACTCCTAG

Additional PACE primers segregating in GVA-H-20-1080 associated with *Autoflower1* are listed in [Supplementary Table 1](#).

Using a delta-allele frequency approach, none of the genome reached significance, but the region identified on Chromosome 1 neared significance, and there were non-significant peaks on Chromosomes 8 and X.

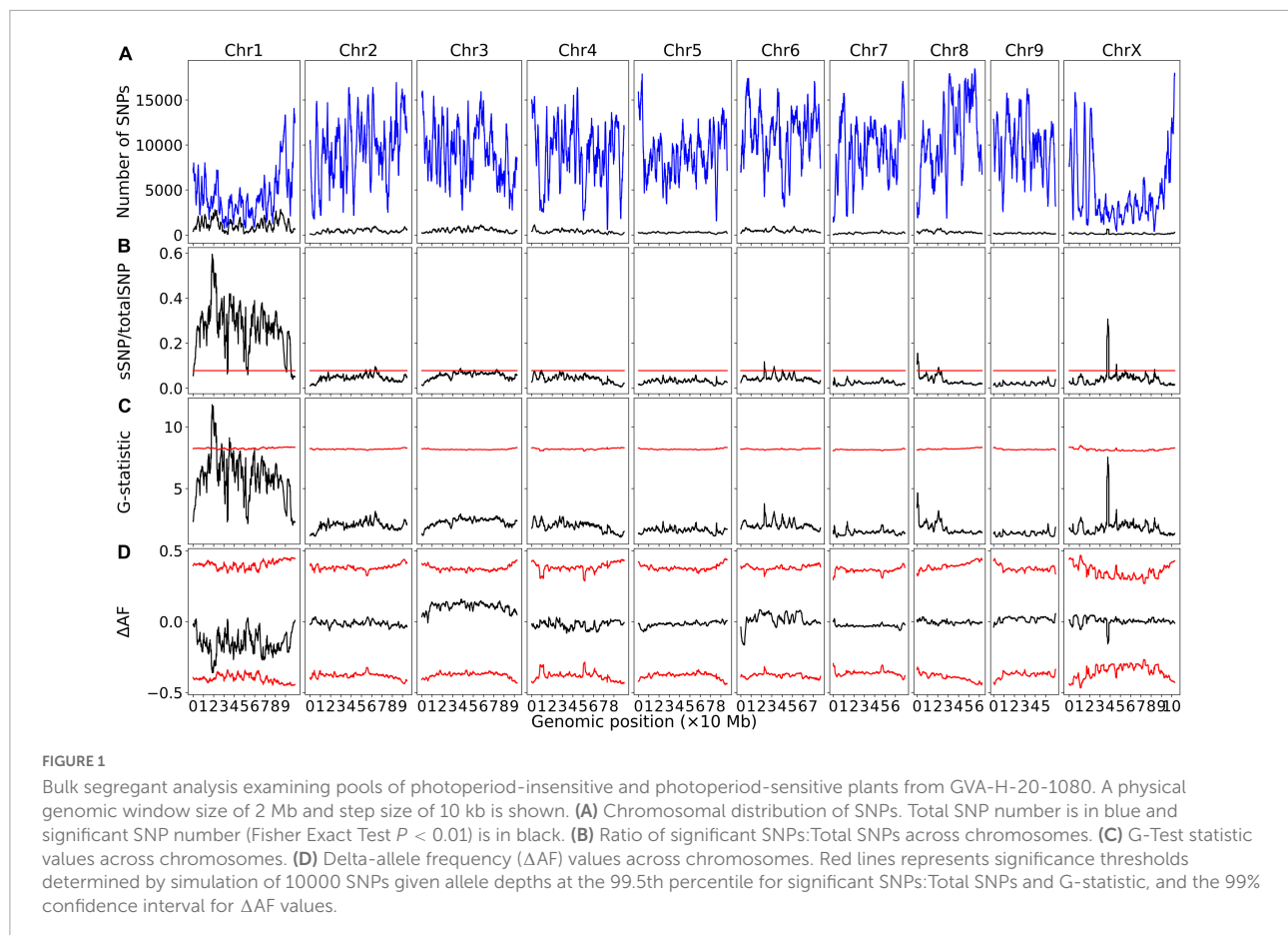
Autoflower1 candidate gene analysis

Within the G-statistic significant region of *Autoflower1* on Chromosome 1 defined by the BSA of GVA-H-20-1080, 237 annotated genes were identified using the NCBI Genome Data Viewer. Of these, 75 were uncharacterized. Candidate genes potentially involved in controlling flowering time include: DOF zinc finger nucleases (LOC115704700, LOC115704742), nuclear transcription factor Y subunit B-1 (NFYB1, LOC115706176), floral homeotic protein APETALA 2 (AP2, LOC115708151), regulator of nonsense transcripts UPF2 (LOC115706264), zinc finger CCCH domain-containing protein 11 (LOC115706080), two-component response regulator-like PRR37 (PRR37, LOC115705128), protein FAR1-RELATED SEQUENCE 5-like

(LOC115703878, LOC115703890), and protein LONG AFTER FAR RED 3 (LOC115705698). The 237 genes within the significant region are detailed in [Supplementary Table 2](#).

Germplasm screening with *Autoflower1* molecular assays

Within the region significantly associated with *Autoflower1*, some SNP alleles that were homozygous in the photoperiod-insensitive bulk and had an allele frequency of ~33% in the photoperiod-sensitive bulk were converted to PACE assays and screened on diverse germplasm ([Table 2](#)). As photoperiod-sensitive plants may be heterozygous for *Autoflower1*, the *Autoflower1* marker assay was considered to be perfect if the homozygous allelic group associated with photoperiod-insensitive plants from the bulk was associated with photoperiod-insensitive plants only, while photoperiod-sensitive plants were either heterozygous or in the alternate homozygous allelic group.



Effect of *Autoflower1* genotype on agronomic performance

Ninety-six individuals of GVA-H-20-1080 grown in the 2021 flowering time field trial were genotyped at *Autoflower1* using the AUTO-2 marker to determine the additive effect of this locus when grown under field conditions. There was a significant effect of the allelic group on flowering date, height, and biomass, with heterozygotes being intermediate with respect to flowering date, height, and wet biomass (Figure 2).

Flowering of diverse germplasm under continuous light

Several populations were grown under continuous light to determine if non-*Autoflower1*/*Autoflower1* cultivars could be induced to flower. There were distinct patterns of flowering time behavior among and within populations (Table 3). Most high-cannabinoid cultivars did not flower under continuous light, except for those homozygous for *Autoflower1*. Notably, plants heterozygous at the *Autoflower1* locus did not flower

under continuous light. In general, cultivated fiber and Chinese cultivars did not terminally flower, while male and female plants in closely related feral populations did produce terminal and axillary (solitary) flowers, respectively. Also flowering under continuous light were the Canadian grain cultivars, ‘Picolo’ and ‘CFX-2’ (Hemp Genetics International, Saskatoon, SK).

Complementation test of photoperiod-insensitive cultivars

The Canadian grain hemp cultivar ‘Picolo’ and a subset of the individuals in the ‘Le Crème’ population both flowered under continuous light, but had contrasting PACE marker calls at the *Autoflower1* locus. A complementation test was performed to determine if there were distinct genes underlying their respective photoperiod insensitivity. All F_1 plants from this cross were induced to flower, although the time to flower was distinct from the parents (Table 4). Homozygous *Autoflower1*/*Autoflower1* female plants flowered 3 weeks earlier than female ‘Picolo’ and female F_1 plants, and male ‘Picolo’ plants flowered 2 weeks earlier than the male F_1 plants. Female ‘Picolo’ and F_1 plants were morphologically similar, while ‘Le

TABLE 2 Genotype group calls for the *Autoflower1* locus by cultivar or population.

Cultivar or population	Source	<i>Autoflower1</i> ¹	n	AUTO-1 ² (18.464 Mb)	AUTO-2 (19.701 Mb)	AUTO-3 (19.731 Mb)	AUTO-4 (19.991 Mb)	AUTO-5 (21.536 Mb)
				CASPL4D1	NFYB1	AP2	PRR37	LOC115703889
'Anka'	UniSeeds Inc.	WT/WT	4	A/A	C/C	C/C	T/T	T/T
Bish Feral	Bish Enterprises	Unknown	8	A/A	C/C	C/C	C/C	T/T
C16	Arcadia	WT/WT	2	A/A	C/C	C/C	T/T	T/T
'CFX-2'	Hemp Genetics Intl.	Unknown	8	A/A	C/C	C/C	C/C	T/T
'Henola'	Intl. Hemp	Unknown	8	A/A	C/C	C/C	T/T	T/T
'Picolo'	Hemp Genetics Intl.	Unknown	16	A/A	C/C	C/C	Seg*	T/T
'Puma'	CN Kenaf and Hemp	WT/WT	4	A/A	C/C	C/C	Seg*	T/T
RN13A	Paul Smith Denver Co.	WT/WT	4	A/A	C/C	C/C	T/T	T/T
RN17	Paul Smith Denver Co.	WT/WT	8	A/A	C/C	C/C	Seg*	T/T
'Si-1'	CN Kenaf and Hemp	WT/WT	19	A/A	C/C	C/C	Seg*	T/T
'Canda'	Parkland Ind.	Unknown	8	A/A	C/C	C/C	C/C	T/T
Missouri Feral	John Fike (40.2, -94.6)	Unknown	8	A/A	C/C	C/C	C/C	T/T
'Nebraska'	Winter Fox Farms	Unknown	8	A/A	C/C	C/C	Seg*	T/T
'NWG-Elite'	New West Genetics	Unknown	8	A/A	C/C	C/C	Seg*	T/T
'T2'	Boring Hemp Co.	WT/WT	8	A/A	C/C	C/C	Seg*	T/T
'USO-31'	UniSeeds Inc.	Unknown	8	A/A	C/C	C/C	C/C	T/T
'CBG Delight'	Flura	Segregating	32	Seg†	Seg†	Seg†	Seg†	Seg*
'H5'	American Hemp Co.	Segregating	32	Seg*	Seg†	Seg†	Seg*	Seg*
'Hempress'	Point3 Farma	Segregating	24	Seg†	Seg†	Seg†	Seg*	Seg†
'Le Crème'	Ventura Seed Co.	Segregating	44	ND	Seg†	Seg†	ND	ND
GVA-H-20-1080	Cornell Hemp	Segregating	184	Seg*	Seg†	Seg†	Seg†	Seg*
'TJ's CBG'	Stem Holdings Agri	Segregating	88	Seg†	Seg†	Seg†	Seg†	Seg*
'Suver Haze'	Oregon CBD	WT/ <i>Autoflower1</i>	8	T/C	T/C	T/C	C/C	C/T
'Umpqua'	Ind. Seed Innovations	WT/ <i>Autoflower1</i>	4	T/C	T/C	T/C	T/T	T/C
AD1010	Phylos Bioscience	<i>Autoflower1</i> / <i>Autoflower1</i>	4	T/T	T/T	T/T	T/T	Seg*

(Continued)

TABLE 2 (Continued)

Cultivar or population	Source	<i>Autoflower1</i> ¹	n	AUTO-1 ²	AUTO-2	AUTO-3	AUTO-4	AUTO-5
				(18.464 Mb)	(19.701 Mb)	(19.731 Mb)	(19.991 Mb)	(21.536 Mb)
				CASPL4D1	NFYB1	AP2	PRR37	LOC115703889
'Alpha Explorer'	Phylos Bioscience	<i>Autoflower1</i> / <i>Autoflower1</i>	4	T/T	T/T	T/T	T/T	Seg*
'Alpha Nebula'	Phylos Bioscience	<i>Autoflower1</i> / <i>Autoflower1</i>	4	T/T	T/T	T/T	T/T	Seg*
'Auto CBD'	Phylos Bioscience	<i>Autoflower1</i> / <i>Autoflower1</i>	4	T/T	T/T	T/T	T/T	C/C
'Auto CBG'	Oregon CBD	<i>Autoflower1</i> / <i>Autoflower1</i>	4	T/T	T/T	T/T	T/T	Seg*
DNCBD	Arcadia Bioscience	<i>Autoflower1</i> / <i>Autoflower1</i>	4	T/T	T/T	T/T	T/T	C/C
'Dr. Chunk'	Kayagene	<i>Autoflower1</i> / <i>Autoflower1</i>	4	T/T	T/T	T/T	T/T	C/C
'Maverick'	Kayagene	<i>Autoflower1</i> / <i>Autoflower1</i>	4	T/T	T/T	T/T	T/T	C/C
'Purple Star'	Atlas Seeds	<i>Autoflower1</i> / <i>Autoflower1</i>	4	T/T	T/T	T/T	T/T	C/C
'Rincon'	Kayagene	<i>Autoflower1</i> / <i>Autoflower1</i>	4	T/T	T/T	T/T	T/T	Seg*
'Sour Citron'	Kayagene	<i>Autoflower1</i> / <i>Autoflower1</i>	4	T/T	T/T	T/T	T/T	C/C
'Sour RNA Seedless' (triploid)	Oregon CBD	<i>Autoflower1</i> / <i>Autoflower1</i>	4	T/T(/T)	T/T(/T)	T/T(/T)	T/T(/T)	C/C(/C)

¹Expected *Autoflower1* locus status based on phenotype and breeding history. Segregating populations all segregating 3:1 photoperiod sensitive: photoperiod insensitive.

²Segt, segregating perfectly; Seg*, segregating imperfectly; ND, not determined.

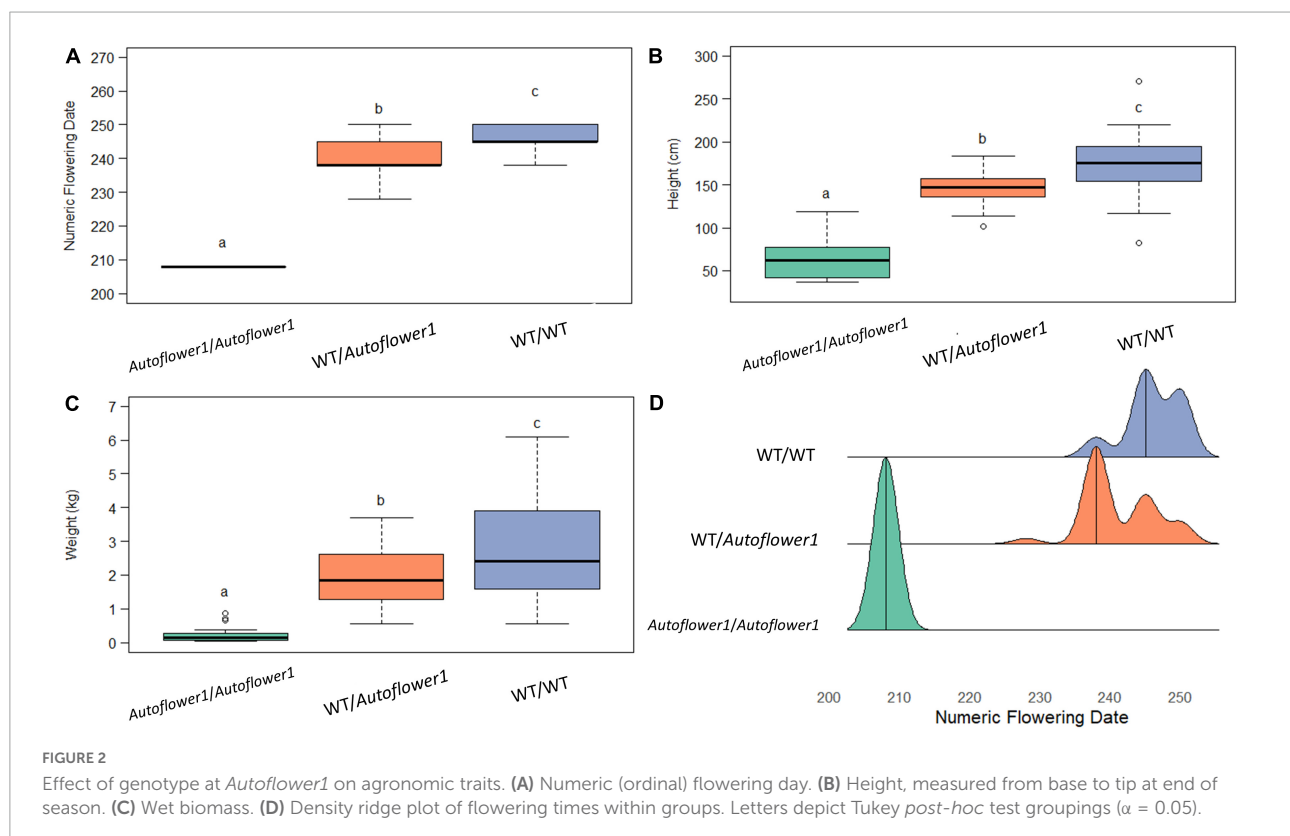


TABLE 3 Time to flower of various *C. sativa* cultivars and populations under continuous light.

Cultivar or population	Males present ¹	Taxonomic group ²	<i>AUTO</i> <i>-FLOWER1</i>	<i>EARLY1</i>	Source	Flowering date ^{1,3}				
						May 10	May 17	May 24	May 31	June 28
'Carmagnola'	+	Fiber/Feral	-/-	-/-	Schiavi Seed	-	-	-	-	-
'Puma'	+	Chinese	-/-	-/-	CN Kenaf and Hemp	-	-	-	-	-
GVA-H-19-1052	-	West Coast	-/-	+/, ±, -/-	Cornell Hemp	-	-	-	-	-
RN16	+	T1/R4	-/-	-/-	Paul Smith Denver Co.	-	-	-	-	-
'Umpqua'	-	West Coast	±	±, -/-	Arcadia Bioscience	-	-	-	-	-
NS52	-	Not tested	±, -/-	±, -/-	Phytonyx	-	-	-	-	-
'Fedora 17'	-	Grain/Dual	-/-	+/, ±, -/-	UniSeeds, Inc.	-	-	-	-	Axillary
'A2R4'	+	Fiber/Feral	-/-	-/-	WinterFox Farms	-	-	-	-	Axillary
'BaOx'	+	BaOx/Otto II	-/-	-/-	Ryes Creek	-	-	-	-	M, axillary F
'Nebraska'	+	Fiber/Feral	-/-	-/-	WinterFox Farms	-	-	-	-	M, axillary F
Missouri Feral	+	Fiber/Feral	-/-	-/-	J. Fike	-	-	-	M only	M, axillary F
GVA-H-20-1080	-	Intercross	+/, ±, -/-	-/-	Cornell Hemp	-	1/4	1/4	1/4	1/4
'Auto CBD'	-	Not tested	+/+	-/-	Phylos Bioscience	-	+	+	+	+
'Auto CBG'	-	Not tested	+/+	-/-	Oregon CBD	-	+	+	+	+
'Socati Auto'	-	Not tested	+/+	-/-	Boring Hemp Co.	-	+	+	+	+
KG9202	-	West Coast	+/+	-/-	Kayagene	-	+	+	+	+
'Anka'	+	Grain/Dual	-/-	-/-	UniSeeds Inc.	-	M only	M only	M, axillary F	M, axillary F
'Henola'	-	Grain/Dual	-/-	+ /+	Bija Hemp	-	Some	+	+	+
'CFX-2'	+	Grain/Dual	-/-	-/-	HGI	M only	+	+	+	+
'Picolo'	+	Grain/Dual	-/-	-/-	HGI	M only	+	+	+	+

¹XY plants present in population; yes (+), no (-); Alleles; non-WT (+), WT (-); Flowers present on date, yes (+), no (-).

²Taxonomical group data described in Carlson et al. (2021).

TABLE 4 Time to flower under long day (16L:8D) lighting.

Cultivar or pedigree	Males (weeks)	Females (weeks)
'AutoCBD', 'Le Crème' (<i>Autoflower1/Autoflower1</i>)	n/a	4 (AutoCBD <i>n</i> = 10, Le Crème (<i>Autoflower1/Autoflower1</i>) <i>n</i> = 3)
'Picolo'	4 (<i>n</i> = 7)	7 (<i>n</i> = 3)
'Le Crème' (<i>Autoflower1/Autoflower1</i>) × 'Picolo' (F ₁)	6 (<i>n</i> = 4)	7 (<i>n</i> = 6)
RN16	None	None

Crème' *Autoflower1* plants were distinct (Figure 3). All plants of a given sex and genotype flowered on the same day.

Segregation for flowering time in 'Umpqua'

The cultivar 'Umpqua' has been grown in Cornell field trials in 2019, 2020, and 2021. In each year, two distinct flowering times were noted (Figure 4). Over the course of 3 years, clear grouping was apparent, with 78 plants total in the early flowering group and 97 plants total in the later flowering group. These data are consistent with a 1:1 segregation of early and late phenotypes ($\chi^2 = 2.063$, $P = 0.15$), characteristic of a backcross involving a major effect gene (designated here as *Early1*) that is heterozygous in one parent and homozygous in the other. Neither phenotype was induced to flower under continuous light (Table 3).

Mapping of *Early1* in 'Umpqua'

Bulk segregant analysis showed clear statistical significance for the *Early1* locus on Chromosome 1 (NC_044371.1), reaching significance using the delta-allele frequency approach (Figure 5). 247814 SNPs on Chromosome 1 were used in this analysis. Examination of the significant SNP data showed that the early flowering 'Umpqua' group was heterozygous at *Early1* while late flowering 'Umpqua' group was not.

Using a G-statistic threshold, most of Chromosome 1, as well as small peaks on Chromosomes 8 and X, were deemed significant.

Early1 candidate gene analysis

For the two 'Umpqua' pools, there were several small peaks that exceeded significance levels for the delta-allele frequency metric on Chromosome 1, spanning the intervals 35.26–36.23,



FIGURE 3
Photoperiod-insensitive *C. sativa* grown under long days (16L:8D) and photographed 85 days after planting. Representative female plants of panel (A) 'Picolo,' (B) 'Le Crème' (*Autoflower1/Autoflower1*), and (C) 'Le Crème' (*Autoflower1/Autoflower1*) × 'Picolo' F₁. Receptive white pistils were present at the apex of plants in panels (A,C) at 85 days after planting, but only brown desiccated pistils were present on plants in panel (B).

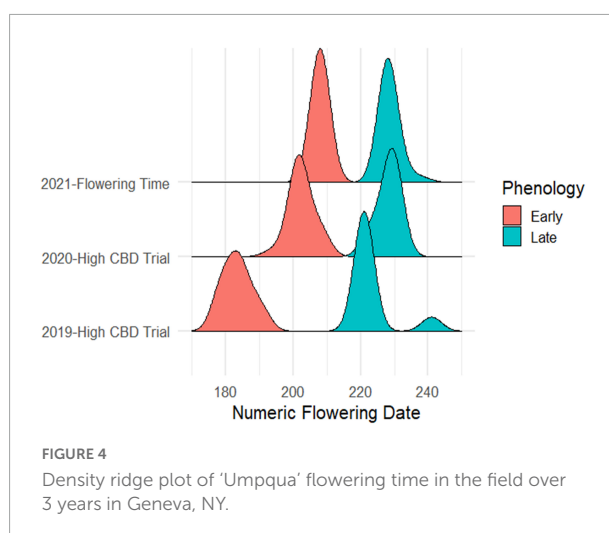
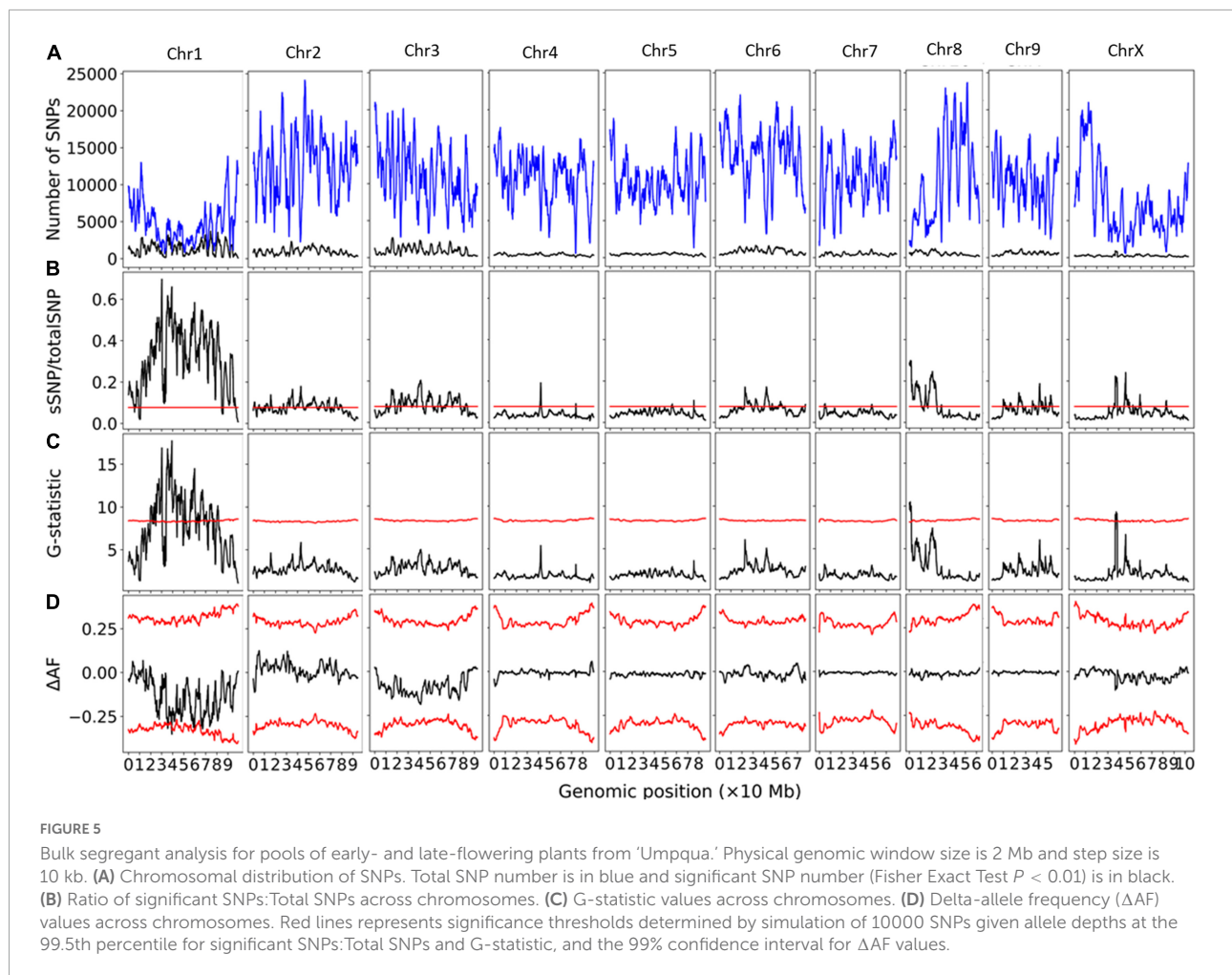


FIGURE 4
Density ridge plot of 'Umpqua' flowering time in the field over 3 years in Geneva, NY.

38.67–39.36, and 59.8–59.9 Mb. A total of 45 genes are annotated within these regions, of which the strongest candidate gene for *Early1* based on molecular function is LOC115705415



(annotated to encode Casein kinase 1-like protein 1), located on Chromosome 1 at 39.265 Mb ([Supplementary Table 3](#)). Polymorphic SNPs that were heterozygous for *Early1* in the early flowering pool were developed into high-throughput PACE marker assays ([Table 1](#)). Genotype assays correlated perfectly with the early- and late-flowering phenotypes of 'Umpqua' across all tested plants ($N = 175$).

Discussion

Autoflower1

Autoflower1 confers photoperiod insensitivity in diverse *C. sativa* germplasm, and segregates in a simple, recessive (Mendelian) manner. Using BSA, we mapped the *Autoflower1* locus derived from KG9202 controlling the photoperiod insensitive phenotype in the GVA-H-20-1080 population to a small region on *C. sativa* Chromosome 1. Polymorphic markers from the Illumina data were used to develop *Autoflower1* molecular assays that accurately reported cultivars marketed as

"Autoflowers," and did not report any photoperiod sensitive plant as photoperiod insensitive. The markers were not effective at predicting photoperiod insensitivity across all germplasm, but this may be due to multiple causes of photoperiod insensitivity.

While *Autoflower1* is recessive with respect to photoperiod insensitivity, plants that were heterozygous for *Autoflower1* flowered approximately 2 weeks earlier than plants that were homozygous for *Autoflower1* under field conditions. This earlier flowering resulted in smaller plants with less total biomass, but may be useful for higher latitudes, as cultivars that are heterozygous for *Autoflower1* can produce very high yields in a shorter growing season ([Stack et al., 2021](#)). Many available cultivars are heterozygous for *Autoflower1*, which may be used as an effective breeding strategy for intellectual property protection. The prevalence of segregating populations marketed as cultivars suggests that some (perhaps unscrupulous or novice) breeders used parents that were heterozygous at *Autoflower1* leading to photoperiod-insensitive plants in the seed population. Populations segregating with $\sim 1/4$ photoperiod-insensitive individuals, such as 'TJ's CBG,' suggest production by a cross between two parents heterozygous for *Autoflower1*, possibly

the self-pollination of a plant heterozygous for *Autoflower1*. As detailed in **Table 2**, several commercially marketed cultivar populations from multiple sources were segregating for *Autoflower1*, which may have resulted in poor cultivar performance for growers due to variation in photoperiod sensitivity.

Further work to identify the taxonomic origin of *Autoflower1* is pertinent. *Autoflower1* is often ascribed in the gray literature as derived from *C. ruderalis*, but the most recent and in-depth genomic studies do not support the existence of this group (Green, 2005; Carlson et al., 2021; Ren et al., 2021). *Autoflower1* would be expected to have evolved either at very high or very low latitudes, where photoperiod insensitivity is evolutionarily advantageous because seasonal variation in daylength is minimal. Further plant collecting expeditions and population genomics analyses would help resolve the evolutionary origin of *Autoflower1* as well as the extent of the genetic and phenotypic diversity of photoperiod sensitivity in *C. sativa*.

Future work to determine the causative gene at *Autoflower1* will allow biotechnological manipulation of the photoperiod sensitivity phenotype and more facile conversion of elite cultivars to and from photoperiod insensitivity. There were strong candidate genes for *Autoflower1* based on annotated predicted molecular function in the significant QTL interval identified for the GVA-H-20-1080 pools. Notably, SNPs near the genes for nuclear transcription factor Y subunit B-1 (NFYB1, LOC115706176) and floral homeotic gene APETALA 2 (AP2, LOC115708151) were in linkage disequilibrium and were perfectly associated with predicted trait phenotype across all individuals tested. These genes have the potential to be causative for the trait, as a Nuclear factor Y gene (DTH8) plays an important repressive role related to photoperiod in rice (Wei et al., 2010) while AP2 homologs are also important flowering time repressors in pepper (Yuan et al., 2021) and Arabidopsis (Yant et al., 2010). Future gene silencing or knockouts of these and other potential candidate genes may lead to identification of the true gene or set of genes responsible for this trait, although a patent is already held covering biotechnological manipulation of genes within this genetic interval (Phylos Bioscience, International Patent WO 2021/097496 A2).

Continuous light

Plants from diverse germplasm had different flowering responses to continuous light. Photoperiod-sensitive high-cannabinoid cultivars and modern European and Chinese fiber cultivars did not flower under continuous light. Fiber cultivars have been selected for their ability to continue to grow vegetatively until late in the season, which maximizes stem biomass yield. Some feral populations, which are closely related to European fiber cultivars (Carlson et al., 2021), displayed male flowering, but not terminal female flowering, perhaps

indicating some selective advantage to photoperiod-insensitive male flowering outside of cultivation. This may also reflect the ancestral genetics of the progenitors of these feral populations, but it is difficult to know the original provenance of their progenitors.

Despite not being reported by the *Autoflower1* markers, Canadian grain cultivars ‘Picolo’ and ‘CFX-2’ flowered readily under continuous light conditions. This could be due to the molecular markers not being polymorphic or effective in these populations, or due to an alternate genetic basis for photoperiod insensitivity. Different genetic mechanisms may be resolved with a complementation test. If the same gene was responsible for photoperiod sensitivity in ‘Picolo’ and *Autoflower1* ‘Le Crème,’ F₁ progeny from an intercross should be uniformly photoperiod-insensitive. Otherwise, other genes, dominance, or epistasis may be involved. The results (**Table 4**) were inconclusive, as all plants flowered under long days, but the timing and architecture of flowering (**Figure 3**) suggests more complex genetic regulation in photoperiod-insensitive plants across broad germplasm.

Early1

Beyond segregation for *Autoflower1*, several elite populations marketed as cultivars have been demonstrated as segregating 1:1 for a major-effect early flowering time phenotype (Stack et al., 2021). The *Early1* locus, which confers an apparent effect size of 2–4 weeks earlier flowering in ‘Umpqua,’ was also mapped to Chromosome 1 using BSA, but to a different location than *Autoflower1*. The apparent different effect in 2019 compared to 2020 and 2021 may have been due to differences in flowering time rating or non-uniform planting dates. Molecular markers for *Early1* were identified and high-throughput assays developed for this locus, which could further aid in development of cultivars with uniform flowering time.

In searching for candidate genes in the confidence interval for the *Early1* locus identified in ‘Umpqua’ populations, only a small portion of Chromosome 1 was found to exceed significance thresholds by the delta-allele frequency method. One possible candidate gene for early flowering within this small interval encodes a Casein kinase 1-like protein 1 (LOC115705415). This gene is homologous to the major flowering time gene *Early flowering 1/Heading date 16* in rice, another short day plant (Hori et al., 2013). Future validation work could involve genetic engineering or genome editing to accomplish gene knockout or gene knock-in to confirm loss or gain of function. The molecular markers and assays for *Early1* presented in this work will be considerably valuable in breeding. Studies to further explore the interactions between these two flowering time loci, *Autoflower1* and *Early1*, will likely lead to a better understanding of the genetics of flowering time and development of stable cultivars with unique flowering times.

As early flowering ‘Umpqua’ plants were heterozygous for both traits, the progeny of an inbred population would be expected to form nine genotypic groups, whose phenotypes would reveal the role of epistasis between these loci.

There were some differences in the statistical outcomes of the BSA for *Autoflower1* in comparison to *Early1* in ‘Umpqua.’ The statistically significant region of *Early1* in ‘Umpqua’ as determined by the G-Test statistic was much larger and broader than that of *Autoflower1*. This is not surprising if this segregation truly is the result of a simple backcross, as recombination occurs only in one parent, rather than in both parents. This reduces the number of recombination events and therefore increases the apparent QTL size. However, analysis using the delta-allele frequency method resulted in a small peak and reliable diagnostic molecular assays were developed for the *Early1* locus.

Curiously, mapping both *Autoflower1* and *Early1* had apparent peaks on Chromosomes 8 and X, reaching statistical significance by the G-statistic threshold in the case of mapping *Early1*. It is unlikely that a 3-locus model explains the observed segregation in flowering time, so this is likely an experimental artifact. The unusual peaks may be due to errors in mapping or genome assembly, with segments on Chromosomes 8 and X being highly homologous to Chromosome 1 leading to inappropriate mapping, or errors in assembly with those significant regions actually residing on Chromosome 1. The CBDRx-cs10 genome used is known to be incomplete and may not be correctly physically ordered (Kovalchuk et al., 2020).

Many *C. sativa* cultivars produced during the rapid expansion of the cannabinoid industry were segregating for flowering time. There is a critical need in the industry to develop uniform and stable cultivars that represent a range of critical photoperiods. Cultivars with known critical photoperiods can be more effectively matched with the latitude of agricultural regions and environmental conditions in controlled environment cultivation. A better understanding of the genetic basis of flowering time in *C. sativa*, coupled with molecular tools to accelerate breeding and selection, will enable the development of new uniform cultivars to meet this need.

Data availability statement

The data presented in this study are deposited in the NCBI SRA repository, accession numbers: SRR20046529-SRR20046532.

Author contributions

JT: conceptualization, methodology, formal analysis, and writing—original draft. GS and CC: conceptualization, resources, and writing—review and editing. LS:

conceptualization, supervision, writing—review and editing, and funding acquisition. All authors contributed to the article and approved the submitted version.

Funding

This work was partially supported by the New York State Department of Agriculture and Markets through grants AC477 and AC483 from Empire State Development Corporation, a grant from the Foundation for Food and Agriculture Research Hemp Research Consortium (NexGen-Hemp-0000000008) in cooperation with Scotts Corporation, and a sponsored research agreement with Pyxus International. The funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Acknowledgments

We are grateful to the technical staff of the Smart Lab, especially Rebecca Wilk, Teagan Zingg, Allison DeSario, Deanna Gentner, Lauren Carlson, Brian Nardone, Michael Quade, Alexander Wares, and McKenzie Schessl. We would like to thank the generous contribution of germplasm by the companies and individuals listed in Table 3.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.991680/full#supplementary-material>

References

- Ban, S., and Xu, K. (2020). Identification of two QTLs associated with high fruit acidity in apple using pooled genome sequencing analysis. *Hortic. Res.* 7:171. doi: 10.1038/s41438-020-00393-y
- Becker, A., Chao, D.-Y., Zhang, X., Salt, D. E., and Baxter, I. (2011). Bulk segregant analysis using single nucleotide polymorphism microarrays. *PLoS One* 6:e15993. doi: 10.1371/journal.pone.0015993
- Blümel, M., Dally, N., and Jung, C. (2015). Flowering time regulation in crops—what did we learn from *Arabidopsis*? *Curr. Opin. Biotechnol.* 32, 121–129. doi: 10.1016/j.copbio.2014.11.023
- Callaway, J. (2002). Hemp as food at high latitudes. *J. Ind. Hemp* 7, 105–117. doi: 10.1300/j237v07n01_09
- Carlson, C. H., Stack, G. M., Jiang, Y., Taşkıran, B., Cala, A. R., Toth, J. A., et al. (2021). Morphometric relationships and their contribution to biomass and cannabinoid yield in hybrids of hemp (*Cannabis sativa*). *J. Exp. Bot.* 72, 7694–7709. doi: 10.1093/jxb/erab346
- Gloss, D. (2015). An overview of products and bias in research. *Neurotherapeutics* 12, 731–734. doi: 10.1007/s13311-015-0370-x
- Grassa, C. J., Weiblen, G. D., Wenger, J. P., Dabney, C., Poplawski, S. G., Timothy Motley, S., et al. (2021). A new Cannabis genome assembly associates elevated cannabidiol (CBD) with hemp introgressed into marijuana. *New Phytol.* 230, 1665–1679. doi: 10.1111/nph.17243
- Green, G. (2005). *The cannabis breeder's bible: The definitive guide to marijuana genetics, cannabis botany and creating strains for the seed market*. San Francisco, CA: Green Candy Press.
- Hori, K., Ogiso-Tanaka, E., Matsubara, K., Yamanouchi, U., Ebana, K., and Yano, M. (2013). Hd16, a gene for casein kinase I, is involved in the control of rice flowering time by modulating the day-length response. *Plant J.* 76, 36–46. doi: 10.1111/tpi.12268
- Jung, C., and Müller, A. E. (2009). Flowering time control and applications in plant breeding. *Trends Plant Sci.* 14, 563–573. doi: 10.1016/j.tplants.2009.07.005
- Kovalchuk, I., Pellino, M., Rigault, P., Van Velzen, R., Ebersbach, J., Ashnest, J. R., et al. (2020). The genomics of Cannabis and its close relatives. *Annu. Rev. Plant Biol.* 71, 713–739. doi: 10.1146/annurev-arplant-081519-040203
- Lubell, J. D., and Brand, M. H. (2018). Foliar sprays of silver thiosulfate produce male flowers on female hemp plants. *Horttechnology* 28, 743–747. doi: 10.21273/HORTTECH04188-18
- Magwene, P. M., Willis, J. H., and Kelly, J. K. (2011). The statistics of bulk segregant analysis using next generation sequencing. *PLoS Comput. Biol.* 7:e1002255. doi: 10.1371/journal.pcbi.1002255
- McPartland, J. M. (2018). Cannabis systematics at the levels of family, genus, and species. *Cannabis Cannabinoid Res.* 3, 203–212. doi: 10.1089/can.2018.0039
- Mouradov, A., Cremer, F., and Coupland, G. (2002). Control of flowering time: Interacting pathways as a basis for diversity. *Plant Cell* 14, S111–S130. doi: 10.1105/tpc.001362
- Murphy, R. L., Klein, R. R., Morishige, D. T., Brady, J. A., Rooney, W. L., Miller, F. R., et al. (2011). Coincident light and clock regulation of pseudoreponse regulator protein 37 (PRR37) controls photoperiodic flowering in sorghum. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16469–16474. doi: 10.1073/pnas.1106212108
- Nordström, K. J., Albani, M. C., James, G. V., Gutjahr, C., Hartwig, B., Turck, F., et al. (2013). Mutation identification by direct comparison of whole-genome sequencing data from mutant and wild-type individuals using k-mers. *Nat. Biotechnol.* 31, 325–330. doi: 10.1038/nbt.2515
- Petit, J., Salentijn, E. M., Paulo, M.-J., Denneboom, C., and Trindade, L. M. (2020). Genetic architecture of flowering time and sex determination in hemp (*Cannabis sativa* L.): A genome-wide association study. *Front. Plant Sci.* 11:1704. doi: 10.3389/fpls.2020.569958
- Ren, G., Zhang, X., Li, Y., Ridout, K., Serrano-Serrano, M. L., Yang, Y., et al. (2021). Large-scale whole-genome resequencing unravels the domestication history of *Cannabis sativa*. *Sci. Adv.* 7:eabg2286. doi: 10.1126/sciadv.abg2286
- Salentijn, E. M., Petit, J., and Trindade, L. M. (2019). The complex interactions between flowering behavior and fiber quality in hemp. *Front. Plant Sci.* 10:614. doi: 10.3389/fpls.2019.00614
- Segawa, T., Nishiyama, C., Tamiru-Oli, M., Sugihara, Y., Abe, A., Sone, H., et al. (2021). Sat-BSA: An NGS-based method using local de novo assembly of long reads for rapid identification of genomic structural variations associated with agronomic traits. *Breed. Sci.* 71:20148. doi: 10.1270/jsbbs.20148
- Song, J., Li, Z., Liu, Z., Guo, Y., and Qiu, L.-J. (2017). Next-generation sequencing from bulked-segregant analysis accelerates the simultaneous identification of two qualitative genes in soybean. *Front. Plant Sci.* 8:919. doi: 10.3389/fpls.2017.00919
- Stack, G. M., Toth, J. A., Carlson, C. H., Cala, A. R., Marrero-González, M. I., Wilk, R. L., et al. (2021). Season-long characterization of high-cannabinoid hemp (*Cannabis sativa* L.) reveals variation in cannabinoid accumulation, flowering time, and disease resistance. *GCB Bioenergy* 13, 546–561. doi: 10.1111/gcbb.12793
- Toth, J. A., Smart, L. B., Smart, C. D., Stack, G. M., Carlson, C. H., Philippe, G., et al. (2020). Limited effect of environmental stress on cannabinoid profiles in high-cannabidiol hemp (*Cannabis sativa* L.). *GCB Bioenergy* 13, 1666–1674. doi: 10.1111/gcbb.12880
- Toth, J. A., Stack, G. M., Cala, A. R., Carlson, C. H., Wilk, R. L., Crawford, J. L., et al. (2011). Development and validation of genetic markers for sex and cannabinoid chemotype in *Cannabis sativa* L. *GCB Bioenergy* 12, 213–222. doi: 10.1111/gcbb.12667
- Toth, J., Pandurangan, S., Burt, A., Mitchell Fetch, J., and Kumar, S. (2018). Marker-assisted breeding of hexaploid spring wheat in the Canadian prairies. *Can. J. Plant Sci.* 99, 111–127. doi: 10.1139/cjps-2018-0183
- Van Bakel, H., Stout, J. M., Cote, A. G., Tallon, C. M., Sharpe, A. G., Hughes, T. R., et al. (2011). The draft genome and transcriptome of *Cannabis sativa*. *Genome Biol.* 12:R102. doi: 10.1186/gb-2011-12-10-r102
- Wei, X., Xu, J., Guo, H., Jiang, L., Chen, S., Yu, C., et al. (2010). DTH8 suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiol.* 153, 1747–1758. doi: 10.1104/pp.110.156943
- Welling, M. T., Liu, L., Kretzschmar, T., Mauleon, R., Ansari, O., and King, G. J. (2020). An extreme-phenotype genome-wide association study identifies candidate cannabinoid pathway genes in Cannabis. *Sci. Rep.* 10, 1–14. doi: 10.1038/s41598-020-75271-7
- Yang, R., Berthold, E., McCurdy, C. R., da Silva Benevenuto, S., Brym, Z. T., and Freeman, J. H. (2020). Development of cannabinoids in flowers of industrial hemp (*Cannabis sativa* L.)—a pilot study. *J. Agric. Food Chem.* 68, 6058–6064. doi: 10.1021/acs.jafc.0c01211
- Yant, L., Mathieu, J., Dinh, T. T., Ott, F., Lanz, C., Wollmann, H., et al. (2010). Orchestration of the floral transition and floral development in *Arabidopsis* by the bifunctional transcription factor APETALA2. *Plant Cell* 22, 2156–2170. doi: 10.1105/tpc.110.075606
- Yuan, X., Fang, R., Zhou, K., Huang, Y., Lei, G., Wang, X., et al. (2021). The APETALA2 homolog CaFFN regulates flowering time in pepper. *Hortic. Res.* 8:208. doi: 10.1038/s41438-021-00643-7
- Zhang, G.-L., Chen, L.-Y., Xiao, G.-Y., Xiao, Y.-H., Chen, X.-B., and Zhang, S.-T. (2009). Bulk segregant analysis to detect QTL related to heat tolerance in rice (*Oryza sativa* L.) using SSR markers. *Agric. Sci. China* 8, 482–487. doi: 10.1016/S1671-2927(08)60235-7
- Zhang, J., and Panthee, D. R. (2020). PyBSASeq: A simple and effective algorithm for bulked segregant analysis with whole-genome sequencing data. *BMC Bioinform.* 21:99. doi: 10.1186/s12859-020-3435-8
- Zhang, M., Anderson, S. L., Brym, Z. T., and Pearson, B. J. (2021). Photoperiodic flowering response of essential oil, grain, and fiber hemp (*Cannabis sativa* L.) cultivars. *Front. Plant Sci.* 12:694153. doi: 10.3389/fpls.2021.694153