Facultative winter accession of *Camelina sativa* (L. Crantz) with early maturity contributes to understanding of the role of *FLOWERING LOCUS C* in camelina flowering

Matthew A. Ott[1*], Ratan Chopra[2], Katherine Frels[3], Anthony Brusa[1], M. David Marks[4], James A. Anderson[1]

[1]Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, MN 55108

[2]CoverCress Inc., Creve Coeur, MO 63132

[3]Department of Agronomy and Horticulture, University of Nebraska, Lincoln, NE 68583

[4]Department of Plant and Microbial Biology, University of Minnesota, St. Paul, MN 55108

[*]Corresponding author: (ottxx142@umn.edu)

Abstract

Camelina is being developed as a winter oilseed cover crop. Early flowering and maturity are desired traits in camelina to allow for relay planting or seeding of a summer annual following camelina harvest. Here we report that while all winter biotype accessions of camelina have a functional allele of *FLOWERING LOCUS C (FLC)* on chromosome 20, there are also at least 20 previously characterized spring biotype accessions that have a functional *FLC* allele at this locus. We observed this by analyzing 76 of 222 camelina accessions that were resequenced by Li et al., (2020) as well as 11 additional accessions resequenced for this analysis. This discovery will inform marker assisted selection efforts that are underway to increase genetic variation in the genetically narrow base of winter camelina germplasm. Furthermore, we optimized a KASP genotyping approach that effectively differentiates winter camelina accessions from the spring accessions based upon a SNV in the *FLC* gene on chromosome 20. We also identified a facultative winter biotype accession of camelina (PI650163-1) that has demonstrated two years of winter-hardiness and has flowered at least a week earlier than the common winter accession, 'Joelle'. A cross between *Camelina sativa* accessions PI 650168 (n=19) and PI 650163-1 (n=20) produced viable progeny to the current F_3 generation. Early maturing winter-hardy camelina will reduce stress on the subsequent soybean crop and improve cropping system yields. A principal components analysis of genetic diversity showed that three replications of this facultative winter accession clustered near the group of winter accessions.

Introduction

There is unprecedented global demand for biofuels and feed byproducts as consumers seek out fossil fuel alternatives to mitigate climate change and energy price volatility. When planted as a winter annual oilseed cover crop, *Camelina sativa* (L. Crantz) may be able to ease some of the supply shortage issues for both the energy and feed sectors. The energy producer Exxon Mobil has partnered with Global Clean Energy to purchase 220 million gallons of camelina oil in the western US (Sustainable Oils Camelina (susoils.com) 2022). Planting camelina as an oilseed cover crop also has been shown to provide

36 important ecosystem services, such as reducing nitrate leaching and providing supplemental nutrition
37 for pollinators, further promoting the crop as a potential solution to several complex challenges (Weyers
38 et al., 2019; Eberle et al., 2015).

39 *Camelina sativa* has been shown to be descended from *Camelina macrocarpa*, with a region of origin
40 spanning the Mediterranean and Eastern Europe (Brock et al., 2018; Chaudhary et al., 2020). It remains
41 unclear exactly when *Camelina sativa* began to be actively cultivated rather than being tolerated as a
42 weed in cultivated flax (Hovsepyan and Willcox, 2008). Two chromosomal states exist for allohexaploid
43 *Camelina sativa* – either n=6+7+7 (20) or n=6+6+7 (19) (Hotton et al., 2020). The genome size was
44 estimated to be 785 Mb, which is relatively small for a hexaploid species (Kagale et al., 2014).

45 In the *Camelina sativa* germplasm within the US's National Genetic Resources Program, only eight of the
46 41 accessions exhibited the winter biotype, while the rest exhibited the spring biotype which is not
47 winter-hardy (Hotton et al., 2020). This leaves plant breeders focused on winter types for their value as
48 a winter cover crop with a narrow genetic base in this self-pollinated species. To increase genetic
49 diversity for key traits in winter camelina, one strategy would be to cross-pollinate winter and spring
50 types, and either select for winter types in the field if feasible, or backcross as necessary to recover the
51 winter growth habit while introducing impactful alleles. Identifying genetic markers that can predict
52 growth habit in *Camelina sativa* would save time otherwise spent waiting for populations to
53 phenotypically segregate for this trait before selection and desirable allele stacking could take place in a
54 breeding program. The Arabidopsis orthologous gene, *FLOWERING LOCUS C* has been previously shown
55 to effectively differentiate between two spring accessions of camelina (DH55 and CO46) and one winter
56 accession (Ames33292, 'Joelle') using qRT-PCR (Anderson et al., 2018). We have optimized a genotyping
57 approach that differentiates winter from spring camelina accessions based upon *FLC* on chromosome 20
58 using DNA-based Kompetitive Allele Specific PCR (KASP). We also resequenced 11 camelina accessions
59 within our collection and downloaded an additional 76 resequenced accessions from data on NCBI

provided by Li et al., (2020) for further analysis of growth habit and the *FLC* gene across a broad spectrum of camelina accessions.

Materials and Methods

Field and indoor screening of *Camelina sativa* germplasm.

In 2015 we requested 420 accessions of *Camelina sativa* from the USDA-National Plant Germplasm System, Plant Genetic Resources of Canada, the Plant Breeding and Acclimatization Institute (IHAR) based in Poland, and Dr. Johann Vollmann at the Universität für Bodenkultur Wien (BOKU), Vienna, Austria for agronomic characterization and winter hardiness evaluation. Each accession was planted in a single, 1.5m row on September 1st, 2015 using a Wintersteiger Seedmech TRM 2200 plotmatic four-row cone planter with a 38 cm row spacing. Approximately 100 seeds were planted per row. On November 25th, 2015, each row was scored for the presence of reproductive stems or "bolting". Accessions were scored as bolting or slow bolting.

Accessions that were scored as slow bolting were planted in triplicate in a randomized complete block design without vernalization in the greenhouse in St. Paul, MN on February 19th, 2016. The greenhouse temperature was maintained near 20 °C and was supplemented with halogen lighting to ensure 16h light was available each day. The height of each plant was measured 34 days after planting.

On September 13th, 2017 we planted a single four-row plot of each *Camelina sativa* accession previously identified as winter types in addition to a single four-row plot USDA-NPGS accession PI650163 (origin – Former Soviet Union) in the field in Rosemount, in addition to multiple plots of the check, Ames 33292 ('Joelle') accession. This experiment was planted as a seed increase. The seeding rate was 0.4 g per plot and a SRES 4-row plot planter was used with a 38 cm row spacing. The seeding depth was set to 1.27 cm, however, the actual seeding depth varied and was generally deeper due to fluffy soil conditions.

82 On October 18th, 2018 a selected camelina line from accession PI 650163, referred to as PI 650163-1 and
83 the winter camelina check accession, Ames 33292 ('Joelle'), were each planted in a 1.5m row in the field
84 in St. Paul. This selection was identified in the aforementioned Rosemount, MN field in May of 2018 as
85 the earliest flowering plant present. The seeding rate was 0.1 g per row, the seeding depth was 1.27 cm
86 and a SRES 4-row plot planter was used with a 38 cm row spacing.

87 On September 9th, 2019 camelina accessions PI 650163-1 and Ames 33292 were planted in duplicate in a
88 St. Paul field in a completely randomized design to quantify differences in flowering time between these
89 accessions. This site-year was planted by hand at a seeding rate of 0.1 g per 1.5m row and a target
90 seeding depth of 1.27 cm.

91 **Resequenced Camelina accessions**

92 Eleven camelina accessions selected for either winter hardiness, seed protein content, or early maturity
93 were resquenced at the University of Minnesota (UMN) and are listed in Table 1., along with the 76
94 camelina accessions that were resequenced by Li et al., (2020) at Montana State University (MSU). The
95 11 UMN resquenced accessions were sequenced in three separate runs by the University of Minnesota
96 Genomics Center with an Illumina NovaSeq 6000 platform on S1, S1, and S4 flow cells and yielded
97 approximately 38x, 90x, and 80x coverage, respectively.

98 **Variant Calling Pipeline**

99 For the Li et al., (2020) resequenced accessions, interleaved fastq.gz sequence files were downloaded
100 from NCBI using Google Cloud Platform. Each interleaved fastq file was de-interleaved into two separate
101 fastq files using a script provided by Daniel Standage on a Biostars forum post
102 (https://www.biostars.org/p/141256/). Each of these files contained one of the two paired end read
103 sequence results. These de-interleaved files were used for aligning to the *Camelina sativa* reference
104 genome (Kagale et al., 2014) downloaded from the Crucifer Genome Initiative (developed by

105 researchers in Dr. Isobel Parkin's lab as a collaboration between Agriculture and Agri-Food Canada and
106 the Global Food Security Institute, http://cruciferseq.ca/). UMN data was released from the University
107 of Minnesota Genomics Center as zipped, de-interleaved files. All sequence data was aligned using the
108 Burrows-Wheeler transformation aligner (version 0.7.17-r1188, Li and Durbin, 2009)(bwa-mem), and
109 Samtools (version 1.14, Danecek et al., 2021) was used for sorting and compressing the elements of
110 each file. Variants were called using BCFtools (version 1.9, Danecek et al., 2021) and the utilities mpileup
111 (using options -Ou -f) and call (using -Ou -mv).

112 **KASP genotyping to distinguish functional from subfunctional alleles of FLC on chromosome 20**

113 KASP primers were designed as a collaboration between us and Dr. Nisha Jain (3CR Bioscience, Enfield,
114 England) and ordered from Integrated DNA Technologies, Inc (Table 3). Tm and binding site affinity were
115 estimated through SnapGene Viewer. The primer mix included 3.5 µL of the HEX/Y primer (fluorescence
116 533-580), 4.5 µL of the FAM/X primer (fluorescence 465-510), 13 µL of the common primer, and 20 µL of
117 nuclease free water. The PCR master mix used was 3CR Bioscience's PACE2.0 Genotyping Master Mix 2x
118 low ROX (003-0006), and was combined with the primer mix at the following rate per reaction: 0.15 µL
119 primer mix, 5 µL PACE2.0 Master Mix, and 5 µL nuclease free water. KASP genotyping PCR settings on a
120 Roche LightCycler® 480 were:

121 1. 95 °C enzyme activation for 15 minutes

122 2. A touchdown temperature cycle starting with an annealing temperature of 59°C for one minute, then
123 a melting temperature of 95°C for 10 seconds, with a step down of the annealing temperature of 2.2°C
124 until an annealing temperature of 51°C was reached.

125 3. An annealing temperature of 50 °C was used for the remainder of the PCR – all other temperatures
126 and durations remained the same.

127 **Bulked segregant analysis for flowering time**

128 From an F_2 population of 75 plants segregating for flowering time from a cross of PI 650163-1 and Ames
129 33292 (Joelle), seven early flowering plants and eight late flowering plants were selected to form DNA
130 pools for bulked segregant analysis. DNA was first isolated from each plant individually using a Qiagen
131 DNeasy® Plant Mini Kit, then an equal amount of DNA from each early flowering plant (180 ng) was
132 pooled and likewise for each late flowering plant (250 ng). This DNA was submitted to the University of
133 Minnesota Genomics Center for whole genome sequencing using NovaSeq on a S4 flow cell to achieve a
134 target coverage of 80x.

135 The whole genome sequence zipped fastq files were first aligned to the *Camelina sativa* reference
136 genome using the Burrows-Wheeler transformation aligner (version 0.7.17-r1188, Li and Durbin,
137 2009)(bwa-mem), and Samtools (version 1.14, Danecek et al. 2021) was used for sorting and
138 compressing the elements of each file. Variants were called using BCFtools (version 1.9, Danecek et al.,
139 2021) and the utilities mpileup (using options -Ou -f) and call (using -Ou -mv). The resulting VCF file was
140 split into two smaller files using Notepad++ version 8.3.3, then formatted in Microsoft Excel 365 Version
141 2108 using the Data tab and the text to columns function to conform to the format required by the R
142 package QTLseqr (Mansfeld and Grumet, 2018, based on the G prime statistic developed by Magwene et
143 al., 2011), and finally merged back into one file in Notepad++ and saved as a CSV file.
144 Variants were filtered using the QTLseqr package and the following parameters: refAlleleFreq = 0.20,
145 minTotalDepth = 100, maxTotalDepth = 400, minSampleDepth = 40, minGQ = 99. The Gprime analysis in
146 QTLseqr was then used with the following parameters: windowSize = 1e6, outlierFilter = "deltaSNP",
147 filterThreshold = 0.1, where the filter threshold represents a tolerance value for filtering out relatively
148 low confidence QTL. Finally, the plotQTLstats function was used with the following parameters: var =
149 "negLog10Pval", plotThreshold = TRUE, q = 0.01, where q represents the false discovery rate (FDR).

Principal components analysis to estimate genetic relatedness of *Camelina sativa* accessions

Previously generated variant call format files were merged into a single variant call format file using BCFtools 1.15, specifically with the following code found from the online resource https://eriqande.github.io/eca-bioinf-handbook/basic-handling-of-vcf-files.html and listed here:

```
bcftools index Acsn_x.vcf.gz

bcftools index Acsn_n.vcf.gz

bcftools merge -Oz Acsn_x.vcf.gz Acsn_n.vcf.gz > Acsn_x_n.vcf.gz.
```

After all vcf files had been merged, variants were randomly sampled using vcflib in three replications from the merged file to develop three separate files with a manageable size. The first replication randomly sampled 10% of the variants, and the second two replications randomly sampled 20% of the variants. The the following code taken from https://speciationgenomics.github.io/filtering_vcfs/ was used for random variant sampling:

```
bcftools view Acsn_x_n.vcf.gz | vcfrandomsample -r 0.2 >
Acsn_x_n_subset.vcf
```

The subset vcf files were loaded into Tassel 5 and the PCA analysis tab was chosen for each file. The file sampled at 10% was not as consistent in principal component values at the files sampled at 20%, which were nearly identical. To determine how similar principal component values were among the three replications of variant sampling performed, the absolute value of the difference between each value for each accession was first calculated in Microsoft Excel 365 (i.e. |10% PC1-20%PC1-1|, |10%PC1-20%PC1-2|, and |20%PC1-1-20%PC1-2| across each accession). Then, the standard deviation was calculated across each difference from each accession. The standard deviation ranged from 12 to 15 for each principal component (1-3) when the values from the 10% sampling replication were compared to either

173 20% sampling replication (i.e. absolute value of the difference of PC values), but was less than one for
174 each principal component when the two 20% sampling replications were compared. Since the files
175 sampled at 20% were nearly identical, this provided confidence that this sampling rate was sufficient to
176 produce repeatable results. The first file sampled at 20% was used for the results section.

177 ***FLC* pathway analysis of winter camelina accession PI 650143**

178 Since winter camelina accession PI 650143 was observed to have a longer vernalization requirement
179 than other winter camelina accessions (data not shown), it was selected for an analysis of key *FLC*
180 pathway genes, such as *VIL1*. A variant call format (VCF) file comparing PI 650143 and Ames33292
181 (Joelle) was generated as described above. This file was loaded into the Integrated Genome Viewer
182 software (version 2.11.9) along with the *Camelina sativa* reference genome and gff3 gene annotations
183 file obtained from the Crucifer Genome Initiative (developed by researchers in Dr. Isobel Parkin's lab as
184 a collaboration between Agriculture and Agri-Food Canada and the Global Food Security Institute,
185 http://cruciferseq.ca/). The Arabidopsis Information Resource (TAIR) database
186 (https://www.arabidopsis.org) was used to develop a list of genes in the *FLC* pathway, and orthologous
187 genes and associated amino acid sequences were found using the Brassica Database - BRAD (Chen et al.,
188 2022) for *Camelina sativa*. Amino acid sequences were compared among camelina accessions using the
189 online software resource Clustal-Omega multiple sequence alignment tool with the accompanying
190 MView result viewer for data visualization (Madeira et al., 2022).

191 **Cross between *Camelina sativa* accessions with different chromosomal states (n=19 and n=20)**

192 On February 4th, 2020, 62 F_2 progeny of a cross made between camelina accession PI 650168 (n=19) and
193 PI 650163-1 (n=20) were planted into 5.08 cm diameter x 17.8 cm length Deepots™ and placed in two
194 61 x 30.5 x 17.1 cm pot racks the greenhouse. The plants were moved to a 4°C cold room on February
195 27th, 2020, and were placed back into the greenhouse on March 19th 2020. Due to Covid-19 challenges,

196 on March 23rd 2020, the plants were moved to the first author's residence – placed on a south-facing
197 porch with white fluorescent lighting (AgroMax® F54T5HO 6400k fluorescent lamps) and 16h/8h
198 days/nights. The date on which a plant first flowered was recorded until May 14th, 2020 – less than half
199 of the plants had not flowered by this date and were considered late flowering, while most plants
200 succumbed to a pathogen that began to emerge on April 26th, 2020.

201 Eight F_2 plants produced F_3 seed, which was tested for germination by placing 25 seeds on tan 38#
202 germination paper (Anchor Paper Co. St. Paul, MN 55101) cut to 13 x 9.5 cm, then sliding the paper into
203 zipseal bags (Ziploc® sandwich bags 16.5 × 14.9 cm × 1.12 mm, Part No. 315885), and then adding 4 mL
204 of Milli-Q® water. The zipseal bags were then placed in a walk-in growth chamber (Sherer-Dual Jet®,
205 Division of the Kysor Industrial Corporation, Marshall, MI. Model No. CER 812, Serial No. 239-17) with
206 continuous 24h white fluorescent lighting (AgroMax® F54T5HO 6400k fluorescent lamps), a PAR of
207 56 µmol $m^{-2}s^{-1}$, and a temperature of 22 °C (Heatcraft Refrigeration Products LLC, Stone Mountain,
208 Georgia. Model No. LWE075BC).

209 **Results and Discussion**

210 **Spring accessions containing winter *FLC* alleles and vice-versa in *Camelina sativa***

211 As has been shown previously, *FLOWERING LOCUS C* (*FLC*) alleles in Arabidopsis can differentiate the
212 spring versus winter growth habit among camelina accessions (Anderson et al., 2018). The *FLC* gene in
213 Arabidopsis codes for a protein that has been shown to suppress the flowering promoter gene,
214 *FLOWERING LOCUS T* (*FT*) (Samach et al., 2000), and is itself suppressed (methylated) by proteins that
215 are coded by genes activated by exposure to cool temperatures (i.e. 5°C ± 1°C) (Sheldon et al., 2000).
216 The subfunctional allele of this gene in spring type camelina has a frameshift mutation in exon 5
217 (Anderson et al., 2018). As we have looked at these alleles more broadly among a greater number of
218 accessions within the camelina germplasm phenotyped for growth habit and resequenced by Li et al.,

219 (2020) (with nearly identical phenotypes reported here as well in Supplementary Table 1), we have
220 found that there are at least 20 exceptional cases in which a spring accession is homozygous for the
221 winter *FLC* allele on chromosome 20, two cases in which a spring accession is homozygous and three
222 cases in which a spring accession is heterozygous for the winter *FLC* allele on chromosome 8, and two
223 cases in which a winter accession is homozygous for the spring *FLC* allele on chromosome 8 (Table
224 1). For the cases that involved a heterozygous genotype where homozygosity is expected in this self-
225 pollinating species, possible explanations are that either there was a rare case of outcrossing or
226 resequencing sample contamination.

227 **Bulked segregant analysis for flowering time**

228 The objective of this experiment was to identify the primary chromosome specific gene(s) controlling
229 flowering time differences between the facultative winter accession of camelina PI 650163-1 and a
230 winter accession Ames 33292 (Joelle). The QTLseqr R-package analysis identified just one major peak of
231 QTL LOD score values ($-\log_{10}$ of variant p-values) in the bulked segregant analysis experiment (Figure 1).
232 The variant (SNP) with the maximum LOD score (5.6) was at position 4.7 Mb on chromosome 20, though
233 the frameshift inducing INDEL in *FLC* on chromosome 20 (Anderson et al., 2018) had a LOD score of 5.4
234 and is at position 4.2 MB. This result supports our claim that *FLC* on chromosome 20 is the primary
235 regulator of flowering time, rather than *FLC* on chromosome 8 or any other regulatory genes related to
236 flowering time. There may be a few smaller signals of SNP effects on chromosomes 4, 10, and 19, but
237 none were comparable to the signal on chromosome 20. Though the fully functional allele of *FLC* on
238 chromosome 8 is very conserved and highly expressed among winter accessions and the subfunctional
239 allele is conserved among spring accessions (Anderson et al., 2018), it appears that it plays a smaller role
240 in controlling flowering time on the growth habit level in camelina, despite the fact that another report
241 showed it has a significant effect on flowering in spring camelina backgrounds (Li et al., 2020).

242 **KASP genotyping approach distinguished between functional and subfunctional alleles of *FLC* on**
243 **chromosome 20**

244 Currently the only genotyping approach to distinguish functional from subfunctional alleles of the *FLC*
245 gene on chromosome 20 is qRT-PCR, which was developed from camelina accessions we shared through
246 an intermediary (Chao et al., 2019). Many plant breeding programs make use of DNA-based Kompetitive
247 Allele Specific PCR (KASP) genotyping, so we aimed to optimize this genotyping approach for this
248 marker. We found that the KASP common primer we used had 25 binding sites throughout the genome,
249 the X primer had 8 binding sites, and the Y primer had 13 binding sites, according to Snap Gene Viewer's
250 default definition of at least 15 bases annealing. However, many of these sites were not adjacent to
251 each other and thus would not allow for off target amplification to a degree that could confound results.
252 This KASP genotyping approach was effective in distinguishing winter from spring type camelina even
253 without the need of pre-amplification PCR to amplify the region of interest. (Figure 2 with specific
254 camelina accessions used in analysis in Supplementary Table 2).

255 **Discovery of facultative winter camelina accession PI650163**

256 The majority of the accessions within our camelina germplasm had a spring-type growth habit
257 (Supplementary Table 1). Based upon our fall 2015 bolting notes, we conducted a follow-up greenhouse
258 study where we planted each accession/line (20) that was scored as slow bolting into pots in triplicate
259 without a vernalizing cold treatment at any point. Fourteen of these accessions did not bolt, PI 650163
260 repeated the slow bolting phenotype observed in fall 2015, four bolted at a moderate rate, and one
261 bolted very quickly (Figure 3). The latter may have been subjected to a more stressful micro-
262 environment in fall 2015.

263

264

265 **Discovery of winter-hardy, early flowering and maturing camelina accession PI 650163-1**

266 For our fall 2016 winter camelina planting, we only planted verified winter camelina accessions.
267 However, we decided to plant PI650163 in fall 2017 in Rosemount, MN with the hypothesis that it may
268 be able to survive a Minnesota winter due to its facultative winter type growth habit. On May 18th, 2018,
269 a single plant from this plot (PI 650163-1) was selected since it was the first plant flowering in this plot
270 as well as in the entire field. PI 650163-1 flowered nine and 17 days before Ames 33292 ('Joelle') in 2019
271 and 2020 in St. Paul, MN, respectively (Table 4). Physiological maturity was also scored in St. Paul, MN in
272 2020, using the BBCH scale developed by Martinelli and Galasso (2011). PI 650163-1 reached 50%
273 maturity nine days before Ames 33292 (Table 4).

274 In canola, an example of the genetic basis of the facultative winter type is increased expression of a
275 repressor of FLC - the *VENALIZATION INSENSITIVE 3* (*VIN3*) gene, which is typically upregulated after
276 exposure to temperatures near 4°C. In facultative winter type canola, *VIN3* has increased expression
277 even at a temperature of 20°C, far above what typically induces its expression of this gene, and thus
278 baseline *FLC* expression is relatively lower in facultative winter type canola than in winter type (Huang et
279 al., 2021). *Camelina sativa* accession PI 650163 and PI 650163-1 did not have any obvious variants in
280 *VIN3*, but one or more other genes in this pathway may have an allele influencing this phenotype.
281 However, since these facultative winter accessions already have a subfunctional allele of *FLC* on
282 chromosome 20, this floral suppression pathway may be less likely to influence this phenotype.
283 Subfunctional genes in the floral promoter pathway such as *FLOWERING LOCUS T* (*FT*) and *CONSTANS*
284 *(CO)* could also be influencing this phenotype, but again, these accessions did not have any obvious
285 variants in *FT* or in *CO*.

286

287

288 **Genetic relatedness of *Camelina sativa* accessions by principal components analysis of SNP markers**

289 An understanding of the genetic relatedness of agronomically important accessions of camelina may
290 provide new approaches to improve breeding strategies. For the principal component analysis
291 presented here, the first principal component explained 9.9% of the observed genetic variation in SNP
292 markers, adding another principal component increased the cumulative proportion of explained genetic
293 variance to 17.7%, and adding another increased it to 22.9%. Analysis of the first two principal
294 components showed that the camelina accessions with a winter growth habit largely clustered together
295 in quadrants two and three of the principal components plot, the facultative winter accessions clustered
296 in quadrant two, and the spring accessions had representatives in all four quadrants with a
297 concentration in quadrants one and two (Figure 4). The three spring accessions that have a winter *FLC*
298 allele on chromosome 8 (Voronezskij, Vavilov 407, and CS219/PI650153) cluster near the winter
299 accessions. The facultative winter accessions also cluster near the winter accessions. This clustering
300 pattern could possibly be explained by an evolutionary trajectory in which these spring and facultative
301 accessions diverged from winter accessions, or they could simply have common ancestral lineage since
302 they appear to originate from the same general geographic regions (Figure 4). This data on relatedness
303 among camelina accessions with variation in growth habit could be leveraged in breeding programs that
304 have the objective of developing improved winter hardy camelina accessions. Progeny of spring by
305 winter crosses may not require as extensive backcrossing to retain winter hardiness if spring accessions
306 more closely related to winter accessions are chosen as parents.

307 There were two outliers that were located well outside of the winter camelina cluster that appear to be
308 accessions of winter camelina. One of these accessions (PI 650168 or CS234) was resequenced both by
309 us as well as Li et al., (2020), and it was positioned with the other winter accessions according to our
310 data (Figure 4). A third resequencing of this accession may be necessary to reconcile this discrepancy.

311 The other outlier was for accession PI 650143 (CS 209), which appeared to be a unique winter camelina
312 accession, discussed in more detail next.

313 **Uniqueness of winter camelina accession PI 650143 (CS 209)**

314 Of all of the winter camelina accessions evaluated, only PI 650143 did not have a winter allele for *FLC* on
315 chromosome 8 and did have a missense mutation in gene *VERNALIZATION INSENSITIVE 3 LIKE-1* (*VIL1*),
316 coding for an amidic histidine rather than a basic glutamine. Glutamine is also conserved at this position
317 for pennycress (*Thlaspi arvense* L.) as well as for a semi-winter line of *Brassica napus* L. (Brana_ZS_V2.0)
318 that was analyzed from the Brassica Database - BRAD (Chen et al., 2022) (Figure 5). The *FLC* allele on
319 chromosome 8 that is largely conserved among winter accessions of camelina does not appear to
320 substantially influence flowering time in a winter type camelina background, so this may help explain
321 why accession PI 650143 exhibits a winter growth habit while harboring allele conserved among spring
322 types at this locus. A potential subfunctional allele in the vernalization response gene *VIL1* may influence
323 flowering time more substantially and is worth investigating further. This accession also does not cluster
324 with the other winter camelina accessions in the principal components analysis plot (Figure 4).
325 Furthermore, a 21-day vernalization treatment at 4 °C was insufficient to induce flowering, though
326 flowering has been observed after exposure to winter under St. Paul, MN field conditions (data not
327 shown). There have been instances in which an accession thought to belong to the *Camelina sativa*
328 species actually belongs to a closely related species, but this accession has been validated as *Camelina*
329 *sativa*, though it has the less common chromosomal state of n=19 (Hotton et al., 2020). This winter
330 camelina accession will require further study to determine its exact vernalization requirement.

331 **Cross between *Camelina sativa* accessions with different chromosomal states (n=19 and n=20)**

332 The 62 F_2 progeny of the cross made between PI 650168 (n=19) and PI 650163-1 (n=20) (chromosome
333 number determined by Hotton et al., 2020), segregated for flowering time (difference of 38+ days) and

334 eight produced viable F_3 seed, while the rest succumbed to a pathogen that began to emerge on April
335 26th, 2020 or 20 days after the first plant flowered (Supplementary Table 3). Accession PI 650168 has the
336 winter growth habit and PI 650163-1 has a facultative winter growth habit. The large variation in
337 flowering dates is indicative of a population segregating for flowering time, and provides evidence that
338 camelina accessions with different chromosomal states can consistently produce viable offspring at least
339 through the F_2 generation and likely beyond.

340 **Conclusion**

341 While we have gained a much greater understanding of camelina genetics with the extensive whole
342 genome resequencing contribution of Li et al., (2020), our analysis of the sequence data for *FLC* pathway
343 genes has led to additional open questions – primarily, what other gene(s) in the vernalization pathway
344 is/are being suppressed in spring camelina accessions with functional *FLC* chromosome 20 alleles?

345 The genetic relatedness among camelina accessions showed strong clustering by within the winter and
346 facultative winter types. Possible approaches to leverage the relatedness data could be to use it for
347 selecting spring type parents for introducing more diversity into winter camelina breeding programs as
348 well as to select unique parents to breed for increased seed size. Crosses between *Camelina sativa*
349 accessions with different chromosome states (n=19 vs. n=20) consistently produce fertile progeny
350 through at least the F_2 generation and likely beyond, which is useful for leveraging genetic variation
351 present within the germplasm. Unique accessions such as PI 650143 can also be further characterized.

352 The new approach we developed using KASP genotyping to differentiate functional versus subfunctional
353 *FLC* alleles on chromosome 20 is expected to accelerate camelina breeding efforts by making genotyping
354 available on widely used platforms and minimizing costs.

355 The facultative winter camelina accession we identified – PI 650163-1 – appears to hold promise for
356 cropping systems in the Upper Midwest, since it has demonstrated robust winter hardiness in the site

357 years of this study, but also flowers and matures earlier, allowing the primary summer annual crops to
358 receive full light interception sooner. This accession appears to be at a midpoint between spring and
359 winter camelina accessions on a whole genome level, but it remains unclear if its facultative winter
360 nature is controlled by few or many genes.

Table 1. Camelina sativa accessions that were resquenced and analyzed for FLOWERING LOCUS C alleles that could predict growth habit. Resequenced by either the University of Minnesota or by Montana State University. Spring growth habit plants with winter alleles or winter growth habit plants with spring alleles are highlighted in orange.

Accession ID	Accession ID Li et al. (2020)	Accession Name	NCBI Run	Growth Habit	FLC8 Winter Allele	FLC20 Winter Allele
PI 650163	CS229UMN	CSS-CAM37	UMN Resequenced	Facultative winter	het	
PI 650163-1	CS229-1UMN	CSS-CAM37	UMN Resequenced	Facultative winter		
PI 650163	CS229	CSS-CAM37	SRR9993039	Facultative winter		
PI 597833	CS201	163-2073-72	SRR9987775	Spring		
PI 652885	CS235	Ames 29309	SRR10028112	Spring		
PI 652886	CS236	Ames 29310	SRR10028114	Spring		
CN 113755	CS110UMN	Bavaria	UMN Resequenced	Spring		x
CAM 180	CS110	Bavaria	SRR9914681	Spring		x
CAM 61	CS030	Boha1	SRR9722050	Spring		
PI 650144	CS210	Boha2	SRR10873755	Spring		
PI 311735	CS199	Borowska	SRR10873743	Spring		x
CAM 223	CS142	Borowska-ihar	SRR9924268	Spring		
CAM 108	CS063	Bronowska	SRR9722689	Spring		x
IHAR 247026	CS211_IHAR_UMN	BRSCHW 28347	UMN Resequenced	Spring		
PI 650145	CS211	BRSCHW-28347	SRR9990156	Spring		x
PI 650146	CS212	BRSCHW-30021	SRR9990305	Spring		
CAM 63	CS031	CAME1	SRR9722051	Spring		
PI 650140	CS206	Came3	SRR9989930	Spring		
PI 650147	CS213	Came4	SRR9990304	Spring		
PI 650152	CS218	CPS-CAM23	SRR10027314	Spring		
PI 650153	CS219	CPS-CAM10	SRR9990527	Spring	x	
PI 633192	CS202	CR-476/65	SRR9988371	Spring		x
PI 633193	CS203	CR-492/94a	SRR9988367	Spring		x
PI 650161	CS227	CSS_CAM35	SRR9992927	Spring		x
PI 650154	CS220	CSS-CAM25	SRR10027318	Spring		
PI 650156	CS222	CSS-CAM29	SRR9990625	Spring		
PI 650157B	CS223	CSS-CAM30	SRR9990812	Spring		
PI 650159	CS225	CSS-CAM33	SRR9992827	Spring		
PI 650160	CS226	CSS-CAM34	SRR9992839	Spring		
PI 650162	CS228	CSS-CAM36	SRR9993037	Spring		
PI 650164	CS230	CSS-CAM38	SRR10027313	Spring		x
PI 650165	CS231	CSS-CAM7	SRR10001240	Spring		
PI 650166	CS232	CSS-CAM8	SRR10027259	Spring		
CAM 74 or 123	CS039 or 071	Czenstochwska-	SRR9722741	Spring		
Ames 31231	CS191	GE.2011-01	SRR9983193	Spring		
Ames 31232	CS192	GE.2011-05	SRR9983388	Spring		x
PI 650148	CS214	Giessen3	SRR9990306	Spring		
PI 650149	CS215	Giessen4	SRR9990308	Spring		x
PI 633194	CS204	GiessenNr.-3	SRR9988728	Spring		x

CAM 171	CS105	Guillemeteau	SRR9905074	Spring		x
CAM 67	CS034	Hoga1	SRR9722216	Spring		
PI 650150	CS216	Hoga2	SRR9990309	Spring		
PI650167B	CS233	Index Seminum 144	SRR10027319	Spring		
CN 113753	CS107UMN	Irkutskaja-mestnaja	UMN Resequenced	Spring		x
CAM 173	CS107	Irkutskaja-mestnaja	SRR10873422	Spring		x
Vavilov 407	UMN	Karelia	UMN Resequenced	Spring	x	
CAM 7 or 57	CS003 or 027	Kirgizskij	SRR9721948	Spring		
CAM 40	CS020	Lindo	SRR9721882	Spring		
PI 304268	CS195	No. 401	SRR9986031	Spring		
PI 304269	CS196	No. 402	SRR9987777	Spring		
PI 650141	CS207	NU-52279	SRR9989959	Spring		
?	?	Omskij	SRR10873371	Spring		
CAM 25	CS007	Omskij-Mestnyi	SRR9721935	Spring		x
CAM 99	CS059	PRFGL.10	SRR9722541	Spring		
CAM 84	CS047	PRFGL.42	SRR9722430	Spring		
CAM 90	CS053	PRFGL.54	SRR9722466	Spring		
MT102	CS251	Przyth	SRR9721197	Spring		
CAM 165	CS099	STAMM04X05	SRR9733108	Spring		
CAM 196	CS118	STAMM06X10B	SRR9915684	Spring		
CAM 257	CS172	STAMM06X11B	SRR9974986	Spring		
CAM 151	CS085	STAMM07X13B	SRR9727008	Spring		
CAM 157	CS091	STAMM10X12	SRR9730513	Spring		x
MT 5	CS001	Suneson	SRR9721195	Spring		x
PI 650151	CS217	Svalof-AM_CSA217	SRR9990325	Spring		
CAM 68	CS035	Svalof-AM_CSA35	SRR9722218	Spring		
CAM 35	CS015	Ukrajinskij	SRR9721670	Spring		x
PI 258366	CS193	Vniimk-17	SRR9983359	Spring		x
CAM 110	CS064	Volynskaja	SRR9722690	Spring		
PI 258367	CS194	Voronezh-349	SRR9986045	Spring		
?	?	Voronezskij	SRR10873332	Spring	x	
CAM 221	CS141	Voronezskij 3392	SRR10873735	Spring	het	
CAM 174	CS108	Voronezskij 3491	SRR9914678	Spring		x
CAM 120	CS070	Wroclawska	SRR9722738	Spring		x
CAM 78	CS041	Zarja socialisma-Auslese-1	SRR9722315	Spring	het	het
PI 650143	CS209	CS-CROO	SRR9990157	Winter		x
PI 650155	CS221	CSS-CAM27	SRR9990588	Winter	x	x
PI 650158	CS224UMN	CSS-CAM31	UMN Resequenced	Winter	x	x
PI 650168	CS234UMN	NE2006-1	UMN Resequenced	Winter	x	x
PI 650168	CS234	NE2006-1	SRR10028113	Winter	x	x
PI 311736	CS200	Przybrodzka	SRR10001246	Winter	x	x
IHAR 502792	CS200-IHAR-UMN	Przybrodzka	UMN Resequenced	Winter	x	x
CAM 30	CS010	Zarja socialisma1	SRR9721666	Winter	no call	x
CAM 172	CS106	Zarja socialisma2	SRR10873423	Winter	x	x
CN 113691	CS040UMN	Zarja socialisma-Auslese-2	UMN Resequenced	Winter	x	x
CAM 705	CS040	Zarja socialisma-Auslese-2	SRR9722307	Winter	x	x
Ames 33292	UMN	Joelle	UMN Resequenced	Winter	x	x

Table 2. Flowering and maturity visual observations made in the field in St. Paul. Single 1.5 m rows of Camelina were planted on Oct. 18th, 2018 (n=1) and Sept. 9th, 2019 (n=3), respectively.

	2019	2020			
Accession	**First Plant Flowered**	**50% Flowered**	**SE**	**50% Phys. Maturity**	**SE**
Ames 33292 ('Joelle')	25-May	18.5-May	1.5	27-Jun	1
PI 650163-1	16-May	1.5-May	1.5	16-Jun	1

Cs_UMN_KASP_FLC20_C	AATCAAATATCTAACCTTTTCTTTGAGGTT
Cs_UMN_KASP_FLC20_X	***GAAGGTGACCAAGTTCATGCT***TTTATTTGTTACAGACAGAACTAATGTT
Cs_UMN_KASP_FLC20_Y	***GAAGGTCGGAGTCAACGGATT***CTTTTATTTGTTACAGACAGAACTAATGTG

371

372 **Table 3.** Primer design for Kompetitive Allele Specific PCR (KASP) genotyping for winter (allele X) versus
373 spring (allele Y) growth habit in Camelina sativa based upon the the previously identified single basepair
374 INDEL in exon 5 of the *FLOWERING LOCUS C* gene on chromosome 20 (Anderson et al., 2018). The bold,
375 italic DNA sequence at the beginning of the X and Y KASP primers represent the sequences that bind to
376 the FAM and HEX fluorescent dyes, respectively.

377

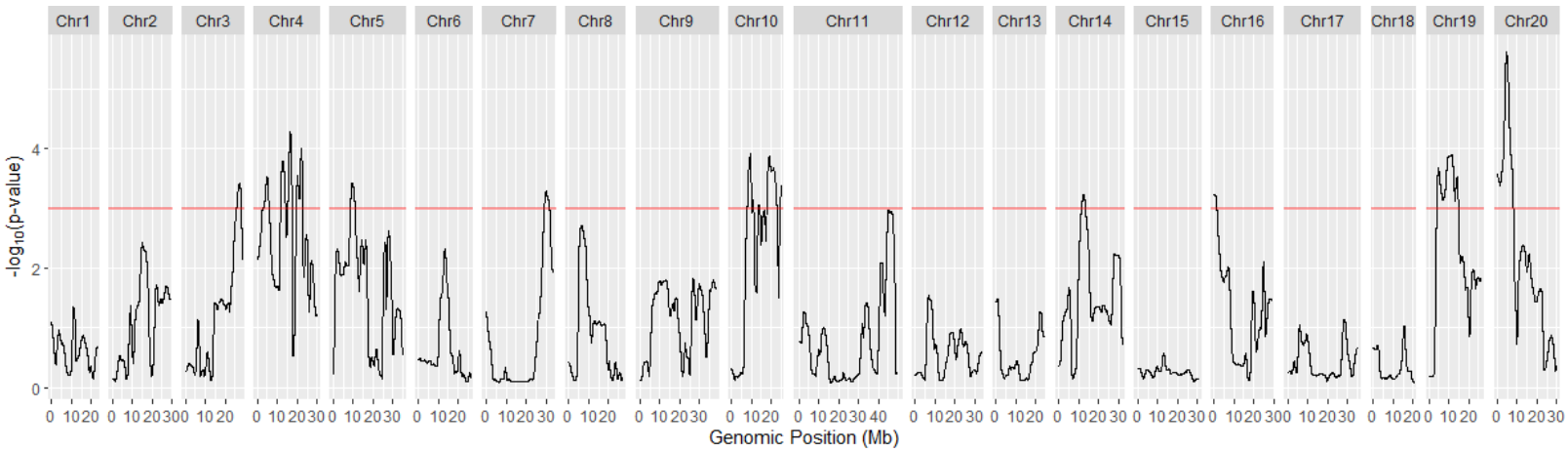
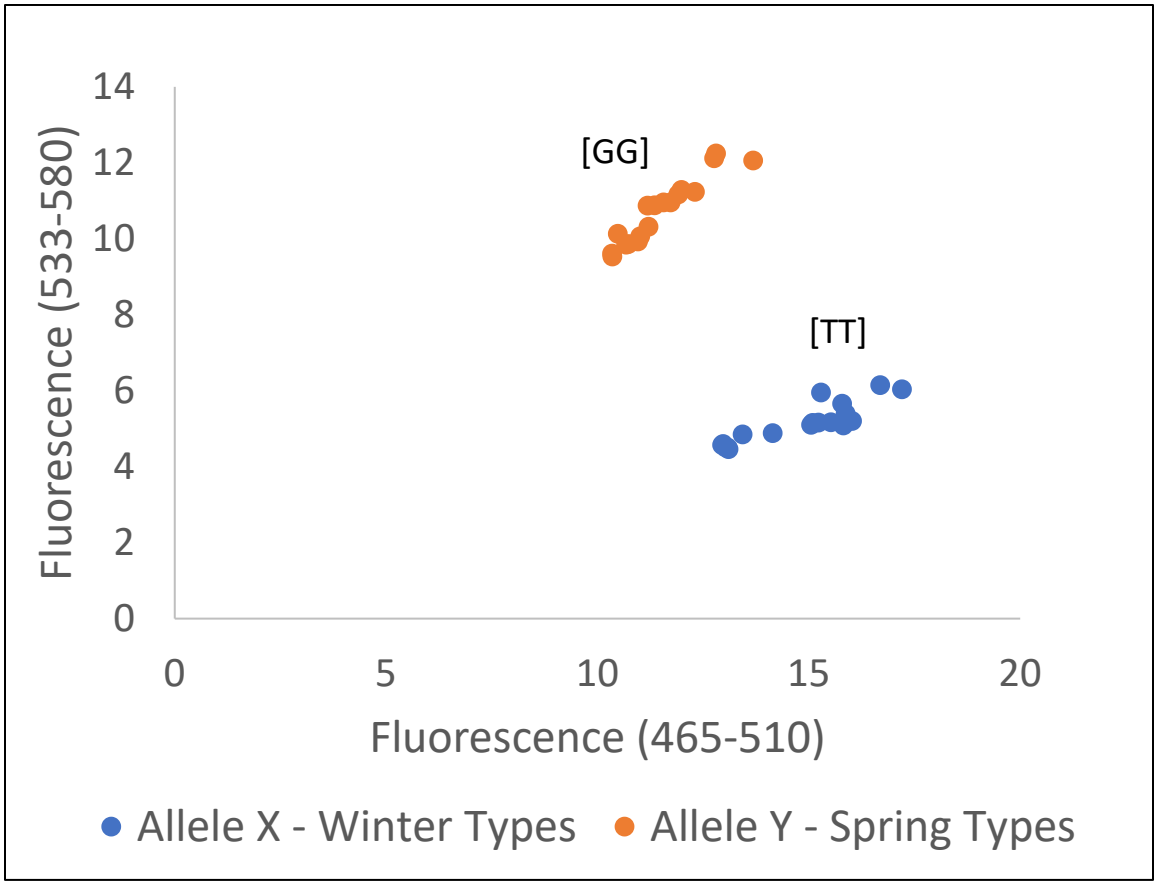


Figure 1. QTL mapping results obtained with QTLseqr and developed using bulked segregant analysis on an F_2 population of a cross between Ames 33292 (Joelle) and PI650163-1. One region on chromosome 20 in *Camelina sativa* L. Crantz is a primary regulator of flowering time. The *FLOWERING LOCUS C* gene is within this region.

383



384

385 **Figure 2.** Kompetitive Allele-Specific PCR (KASP) genotyping results that differentiate winter from spring
386 growth habit accessions of *Camelina sativa* using the previously identified single basepair INDEL in exon
387 5 of the *FLOWERING LOCUS C* gene on chromosome 20 (Anderson et al., 2018). Specific accessions used
388 for this analysis are listed in Supplementary Table 2.

389

390

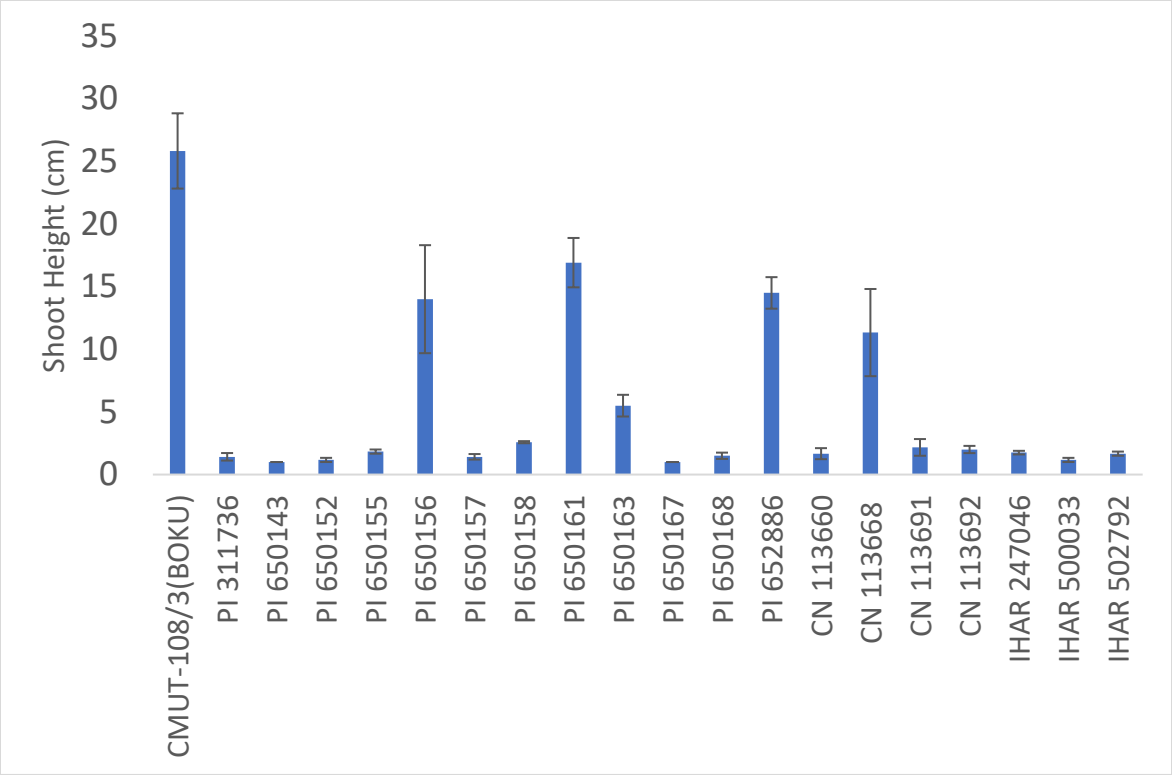


Figure 3. Camelina shoot height observed in a greenhouse 34 days after planting on February, 19th 2016 with no vernalization. Experimental design was a randomized complete block and n=3. The greenhouse temperature was maintained near 20 °C and was supplemented with halogen lighting to ensure 16h light was available each day.

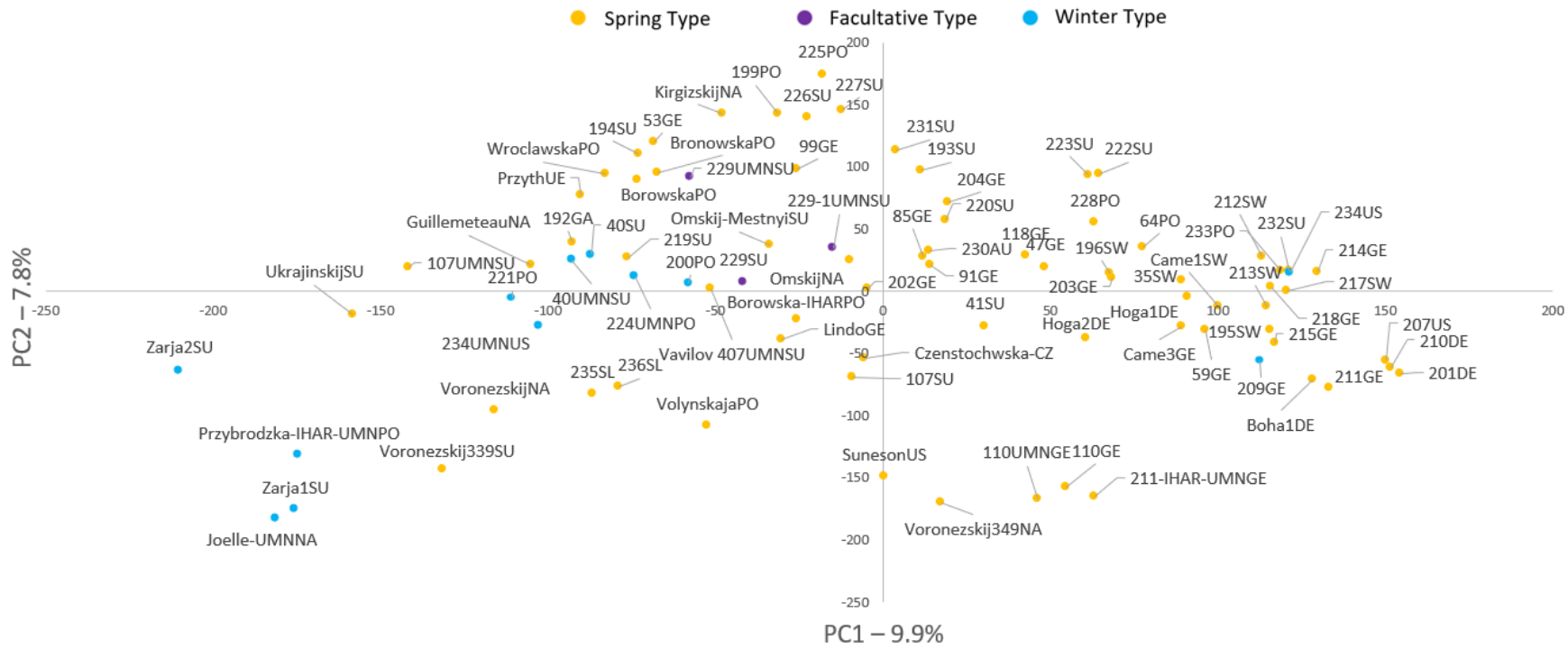


Figure 4. Principal components analysis plot of genetic relatedness among *Camelina sativa* L. Crantz accessions based upon a random sampling of 20% of the genomewide SNP markers. Numeric naming convention established by Li et al., (2020). Suffix two letter country of origin codes are as follows: AU:Austria, CZ: Czech Republic, DE:Denmark, GA:Georgia, GE:Germany, NA:Unknown, PO:Poland, SL: Slovakia, SU: Former Soviet Union, SW: Sweden, UE: Ukraine, US: United States.

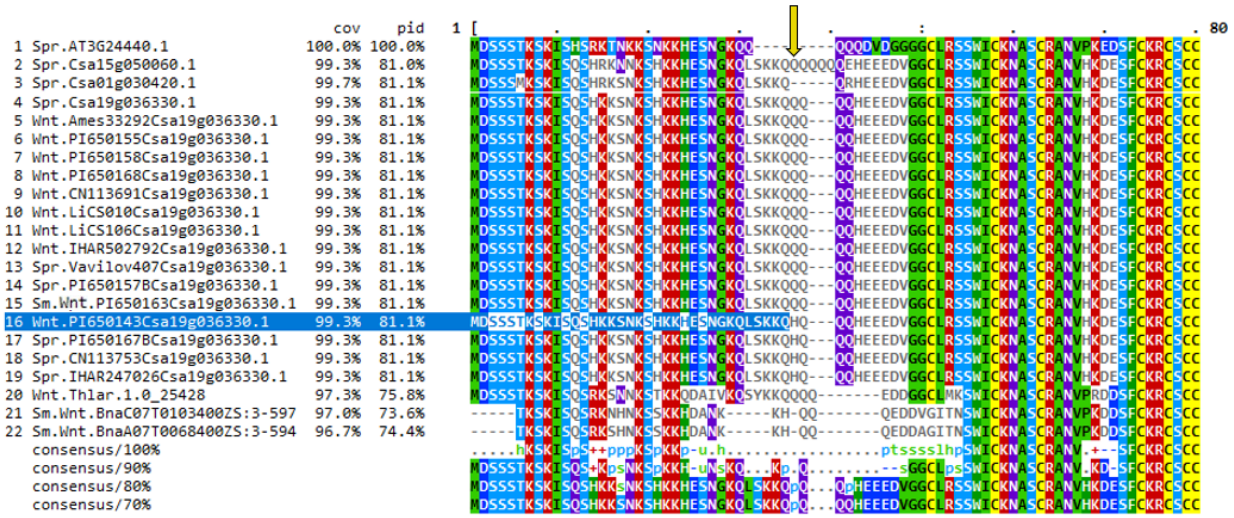


Figure 5. Amino acid sequence alignments for the gene *VERNALIZATION INSENSITIVE 3 LIKE-1* for relevant accessions of camelina (*Camelina sativa* L. Crantz), as well as for a semi-winter canola (*Brassica napus* L.) line (Bna-ZS), winter pennycress (*Thlaspi arvense* L.), and spring Arabidopsis (*Arabidopsis thaliana* L. Heynh).

Acknowledments

The authors would like to thank Dr. Russ Gesch (USDA-ARS, Morris, MN) for providing important agronomic insights into camelina and for providing seed of accession Ames 33292 ('Joelle') before it became available in the NPGS. We would like to thank Dr. Laura Marek – curator of oil seeds crops at North Central Regional Plant Introduction Station (NCRPIS), part of the National Plant Germplasm System (NPGS) – for assistance in establishing a collection of *Camelina sativa* (L. Crantz) at the University of Minnesota. We would also like to thank our lab technician, Serena Gjesvold, and undergraduate student employees, Clair Biel and Henry Parks, for their assistance with lab and field work.

References

Anderson, J. V, Horvath, D.P., Doğramaci, M., Dorn, K.M., Chao, W.S., Watkin, E.E., Hernandez, A.G., Marks, M.D., Gesch, R., 2018. Expression of FLOWERING LOCUS C and a frameshift mutation of this gene on chromosome 20 differentiate a summer and winter annual biotype of Camelina sativa. Plant direct 2, e00060.

Brock, J.R., Donmez, A.A., Beistein, M.A., Olsen, K.M., 2018. Phylogenetics of Camelina Crantz. (Brassicaceae) and insights on the origin of gold-of-pleasure (Camelina sativa). Mol. Phylogenet. Evol. 127, 834–842. https://doi.org/10.1016/j.ympev.2018.06.031

Chao, W.S., Wang, H., Horvath, D.P., Anderson, J. V, 2019. Selection of endogenous reference genes for qRT-PCR analysis in Camelina sativa and identification of FLOWERING LOCUS C allele-specific markers to differentiate summer-and winter-biotypes. Ind. Crops Prod. 129, 495–502.

Chen H, W.T.H.X.C.X.L.R.L.J.W.J.K.G.W.X.. V. 50, I.D. 7 J. 2022, P.D. http://brassicadb. cn/#. A. 10 M. 2022., 2022. Brassicaceae Database (BRAD). Nucleic Acids Res. 50, D1432–D1441.

Crucifer Genome Initiative, n.d. Agriculture and Agri-Food Canada Saskatoon Research and Development Center in collaboration with the Global Institute for Food Security and Dr. Isobel Parkin's lab. [WWW Document].

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H., 2021. Twelve years of SAMtools and BCFtools. Gigascience 10. https://doi.org/10.1093/gigascience/giab008

Eberle, C.A., Thom, M.D., Nemec, K.T., Forcella, F., Lundgren, J.G., Gesch, R.W., Riedell, W.E., Papiernik, S.K., Wagner, A., Peterson, D.H., Eklund, J.J., 2015. Using pennycress, camelina, and canola cash cover crops to provision pollinators. Ind. Crops Prod. 75, 20–25. https://doi.org/10.1016/j.indcrop.2015.06.026

Gesch, R.W., Cermak, S.C., 2011. Sowing date and tillage effects on fall-seeded camelina in the northern corn belt. Agron. J. 103, 980–987.

Hotton, S.K., Kammerzell, M., Chan, R., Hernandez, B.T., Young, H.A., Tobias, C., McKeon, T., Brichta, J., Thomson, N.J., Thomson, J.G., 2020. Phenotypic examination of Camelina sativa (L.) Crantz accessions from the USDA-ARS national genetics resource program. Plants 9, 642.

Hovsepyan, R., Willcox, G., 2008. The earliest finds of cultivated plants in Armenia: evidence from charred remains and crop processing residues in pise from the Neolithic settlements of Aratashen and Aknashen. Veg. Hist. Archaeobot. 17, S63–S71. https://doi.org/10.1007/s00334-008-0158-6

Huang, L.Y., Y. Min, S. Schiessl, X.H. Xiong, H.U. Jan, et al. 2021. Integrative analysis of GWAS and transcriptome to reveal novel loci regulation flowering time in semi-winter rapeseed. Plant Sci. 310. doi: 10.1016/j.plantsci.2021.110980.

Kagale, S., Koh, C.S., Nixon, J., Bollina, V., Clarke, W.E., Tuteja, R., Spillane, C., Robinson, S.J., Links, M.G., Clarke, C., Higgins, E.E., Huebert, T., Sharpe, A.G., Parkin, I.A.P., 2014. The emerging biofuel crop Camelina sativa retains a highly undifferentiated hexaploid genome structure. Nat. Commun. 5. https://doi.org/10.1038/ncomms4706

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. BIOINFORMATICS 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

460 Li, H., Hu, X., Lovell, J.T., Grabowski, P.P., Mamidi, S., Chen, C., Amirebrahimi, M., Kahanda, I., Mumey,
461 B., Barry, K., Kudrna, D., Schmutz, J., Lachowiec, J., Lu, C.F., 2021. Genetic dissection of natural
462 variation in oilseed traits of camelina by whole-genome resequencing and QTL mapping. Plant
463 Genome 14. https://doi.org/10.1002/tpg2.20110

464 Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov,
465 A., Lopez, R., n.d. Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids
466 Res. https://doi.org/10.1093/nar/gkac240

467 Martinelli, T., Galasso, I., 2011. Phenological growth stages of Camelina sativa according to the extended
468 BBCH scale. Ann. Appl. Biol. 158, 87–94.

469 Ott, M.A., Eberle, C.A., Thom, M.D., Archer, D.W., Forcella, F., Gesch, R.W., Wyse, D.L., 2019. Economics
470 and agronomics of relay-cropping pennycress and camelina with soybean in Minnesota. Agron. J.
471 111, 1281–1292.

472 Samach, A., H. Onouchi, S.E. Gold, G.S. Ditta, Z. Schwarz-Sommer, et al. 2000. Distinct roles of
473 CONSTANS target genes in reproductive development of Arabidopsis. Science. 288, 1613–1616.
474 doi: 10.1126/science.288.5471.1613.

475 Sheldon, C.C., D.T. Rouse, E.J. Finnegan, W.J. Peacock, and E.S. Dennis. 2000. The molecular basis of
476 vernalization: The central role of FLOWERING LOCUS C (FLC). Proc. of the Nat. Acad. of Sci. of the
477 USA. 97, 3753–3758. doi: 10.1073/pnas.060023597.

478 Standage, D., 2015. deinterleave fastq file [WWW Document]. Biostars, Bioinforma. Explain.

479 Weyers, S., Thom, M., Forcella, F., Eberle, C., Matthees, H., Gesch, R., Ott, M., Feyereisen, G., Strock, J.,
480 Wyse, D., 2019. Reduced Potential for Nitrogen Loss in Cover Crop-Soybean Relay Systems in a
481 Cold Climate. J. Environ. Qual. 48. https://doi.org/10.2134/jeq2018.09.0350

482 Wittenberg, A., Anderson, J. V, Berti, M.T., 2020. Crop growth and productivity of winter camelina in
483 response to sowing date in the northwestern Corn Belt of the USA. Ind. Crops Prod. 158, 113036.

484