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], Eva Serena Gjesvold[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 75 accessions (67 spring type, one facultative, and seven winter type) that were resequenced by Li et al., (2020) as well as 21 additional accessions 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 the presence of either the functional or subfunctional *FLC* allele on chromosome 20. These analyses identified a facultative winter biotype accession of camelina (PI650163-1, winter hardy with subfunctional chromosome 20 *FLC* allele) that has demonstrated two years of winter-hardiness and has flowered at least a week earlier than the common winter accession, 'Joelle'. A bioinformatics approach to cytotype analysis in camelina also provided more precise categorizing of camelina accessions in the USDA-NPGS germplasm into 2n=38 and 2n=40 cytotypes. Early maturing winter-hardy camelina will reduce stress on a subsequent soybean crop and improve total cropping system yields when camelina and soybean are grown sequentially in the same season on the same land.

1. 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. In the Upper Midwest winter camelina is being developed as a new cover/cash crop that can be planted in the fall and harvested in spring on the same lands used for traditional summer crop such as that undergoing the maize to soybean rotation on much the Midwestern farm acreage. When planted as a winter annual oilseed cover crop, *Camelina sativa* (L. Crantz) should be able to ease some of the supply shortage issues for both the energy and feed sectors. Fall pllanted/spring harvested camelina as an oilseed cover crop also provides important ecosystem services, such as reducing nitrate leaching and providing supplemental nutrition for pollinators, further promoting the crop as a potential solution to several complex challenges (Weyers et al., 2019; Eberle et al., 2015).

Camelina sativa has been shown to be descended from *Camelina macrocarpa*, with a region of origin spanning the Mediterranean and Eastern Europe (Brock et al., 2018; Chaudhary et al., 2020). It remains unclear exactly when *Camelina sativa* began to be actively cultivated rather than being tolerated as a weed in cultivated flax (Brock et al., 2018), but its use as a crop can be documented as far back as the early Iron Age (~1200 BCE) (Zohary et al., 2012). Two cytotypes exist for allohexaploid *Camelina sativa* – either n=6+7+7 (20) or n=6+6+7 (19) (Chaudhary et al., 2020; Hotton et al., 2020; Brock et al., 2022). The genome size was estimated to be 785 Mb, which is relatively small for a hexaploid species (Kagale et al., 2014). Camelina is closely related to Arabidopsis thaliana as both are in Brassicacea

60 lineage group I (Huang et al., 2016). Similar to Arabidopsis, camelina is largely a
61 self-fertilized species. In the *Camelina sativa* germplasm within the US's National
62 Genetic Resources Program, only eight of the 41 accessions exhibited the winter
63 biotype, while the rest exhibited the spring biotype and are not winter-hardy
64 (Hotton et al., 2020). This leaves plant breeders focused on winter types for their
65 value as a winter cover crop with a narrow genetic base. To increase genetic
66 diversity for key traits in winter camelina, one strategy is to cross-pollinate winter
67 and spring types, and either select for winter types in the field if feasible, or
68 backcross as necessary to recover the winter growth habit while introducing
69 impactful alleles. Identifying genetic markers that can predict the winter/spring
70 growth habits in *Camelina sativa* would save time otherwise spent waiting for
71 populations to phenotypically segregate for this trait before selection and desirable
72 allele stacking could take place in a breeding program. In Arabidopsis, a functional
73 *Flowering Locus C* (*FLC*) gene was found to impart a requirement for vernalization
74 ahead of flowering, which the hallmark of a winter type. Similarly in camelina, the
75 presence of functional or subfunctional *FLC* orthologs have been shown to
76 effectively differentiate between two spring accessions of camelina (DH55 and
77 CO46) and one winter accession (Ames33292, 'Joelle') using qRT-PCR (Anderson et
78 al., 2018). These analyses identified a SNP in FLC on chromosome 20 that
79 distinguishes between the spring and winter habits. We have used this information
80 to optimize a genotyping approach that differentiates winter from spring camelina
81 accessions based upon the chromosome 20 FLC SNP using DNA-based Kompetitive
82 Allele Specific PCR (KASP). To expand the our information on the importance of this

SNP in controlling flowering time we resequenced 21 camelina accessions within our collection and downloaded an additional 75 resequenced accessions from data on NCBI provided by Li et al., (2020). While this *FLC* SNP helped differentiate between most spring and winter lines, we identified several spring lines with the winter type *FLC* allele. This information is being used to facilitate the breeding of winter hardy lines that flower and mature early.

2. Materials and Methods

2.1 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 slow bolting 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 uneven soil conditions.

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

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

2.2 Resequenced Camelina accessions

Twenty one camelina accessions were resquenced at the University of Minnesota (UMN) along with the 75 camelina accessions that were resequenced by Li et al., (2020) at Montana State University (MSU). The 21 UMN resquenced accessions were sequenced in

131 four separate runs by the University of Minnesota Genomics Center with an Illumina
132 NovaSeq 6000 platform on S1, S1, S4, S4 flow cells and yielded approximately 38x, 90x, 80x,
133 20x coverage, respectively.

134 **2.3 Variant Calling Pipeline**
135

136 For the Li et al., (2020) resequenced accessions, interleaved fastq.gz sequence files
137 were downloaded from NCBI using Google Cloud Platform. Each interleaved fastq file was
138 de-interleaved into two separate fastq files using a script provided by Daniel Standage on a
139 Biostars forum post (https://www.biostars.org/p/141256/). Each of these files contained
140 one of the two paired end read sequence results. These de-interleaved files were used for
141 aligning to the *Camelina sativa* reference genome (Kagale et al., 2014) downloaded from the
142 Crucifer Genome Initiative (developed by researchers in Dr. Isobel Parkin's lab as a
143 collaboration between Agriculture and Agri-Food Canada and the Global Food Security
144 Institute, http://cruciferseq.ca/). UMN data was released from the University of Minnesota
145 Genomics Center as zipped, de-interleaved files. All sequence data was aligned using the
146 Burrows-Wheeler transformation aligner (version 0.7.17-r1188, Li and Durbin, 2009)(bwa-
147 mem), and Samtools (version 1.14, Danecek et al., 2021) was used for sorting and
148 compressing the elements of each file. Variants were called using BCFtools (version 1.9,
149 Danecek et al., 2021) and the utilities mpileup (using options -Ou -f) and call (using -Ou -
150 mv).

151 **2.4 Alignment of WGS reads for cytotype determination**
152

153 WGS reads from 14 camelina accessions were aligned to the camelina sativa
154 reference genome specifically for determining their cytotype as has been done previously
155 (Chaudhary et al., 2020). This alignment protocol used Bowtie2 version 2.3.4.1 with the

156 sensitive local parameter and the minimum alignment score parameter of 20 + 10 * ln(read
157 length), or specifically in the Bowtie2 language: --score-min G, 20, 10. Then, Bedtools
158 version 2.29.2 was used to count the number of aligned reads in 100 kb windows that had
159 non-zero coverage. This coverage data was then used to generate a Circos plot using
160 TBtools-II version 1.120.

161 **2.5 KASP genotyping to distinguish functional from subfunctional alleles of**
162 **FLC on chromosome 20**
163

164 DNA was isolated from a small leaf piece (~1mm x 2mm leaf sample) using Sigma-
165 Aldrich® Extract-N-AMP DNA extraction (SKU: E7526) and dilution (SKU: D5688) solutions
166 from MilliporeSigma Inc. (St. Louis, MO 63178). KASP primers were designed as a
167 collaboration between us and Dr. Nisha Jain (3CR Bioscience, Enfield, England) and ordered
168 from Integrated DNA Technologies, Inc (Table 1). Tm and binding site affinity were
169 estimated through SnapGene Viewer. The primer mix included 3.5 µL of the HEX/Y primer
170 (fluorescence 533-580), 4.5 µL of the FAM/X primer (fluorescence 465-510), 13 µL of the
171 common primer, and 20 µL of nuclease free water. The PCR master mix used was 3CR
172 Bioscience's PACE2.0 Genotyping Master Mix 2x low ROX (003-0006), and was combined
173 with the primer mix at the following rate per reaction: 0.15 µL primer mix, 5 µL PACE2.0
174 Master Mix, 5 µL nuclease free water and 1 ul of DNA. KASP genotyping PCR settings on a
175 Roche LightCycler® 480 were:
176 1. 95 °C enzyme activation for 15 minutes
177 2. A touchdown temperature cycle starting with an annealing temperature of 59°C for one
178 minute, then a melting temperature of 95°C for 10 seconds, with a step down of the
179 annealing temperature of 2.2°C until an annealing temperature of 51°C was reached.
180 3. An annealing temperature of 50 °C was used for the remainder of the PCR – all other
181 temperatures and durations remained the same.

182 **2.6 Bulked segregant analysis for flowering time**
183

184 From an F_2 population of 75 plants segregating for flowering time from a cross of PI
185 650163-1 and Joelle, seven early flowering plants and eight late flowering plants were
186 selected to form DNA pools for bulked segregant analysis. DNA was first isolated from each
187 plant individually using a Qiagen DNeasy® Plant Mini Kit, then an equal amount of DNA
188 from each early flowering plant (180 ng) was pooled and likewise for each late flowering
189 plant (250 ng). This DNA was submitted to the University of Minnesota Genomics Center for
190 whole genome sequencing using NovaSeq on a S4 flow cell to achieve a target coverage of
191 80x.

192 The whole genome sequence zipped fastq files were first aligned to the *Camelina*
193 *sativa* reference genome using the Burrows-Wheeler transformation aligner (version
194 0.7.17-r1188, Li and Durbin, 2009)(bwa-mem), and Samtools (version 1.14, Danecek et al.
195 2021) was used for sorting and compressing the elements of each file. Variants were called
196 using BCFtools (version 1.9, Danecek et al., 2021) and the utilities mpileup (using options -
197 Ou -f) and call (using -Ou -mv). The resulting VCF file was re-formatted in R using the
198 package 'splitstackshape' version 1.4.8 by splitting the "FORMAT" column for both the late
199 and early flowering pools on the characters of comma and colon to create additional
200 columns that isolated allelic depths and genotype quality scores for each pool conforming to
201 the format required by the R package QTLseqr (Mansfeld and Grumet, 2018, based on the G
202 prime statistic developed by Magwene et al., 2011).

203 Variants were filtered using the QTLseqr package and the following parameters:
204 refAlleleFreq = 0.20, minTotalDepth = 100, maxTotalDepth = 400, minSampleDepth = 40,
205 minGQ = 99. The Gprime analysis in QTLseqr was then used with the following parameters:
206 windowSize = 1e6, outlierFilter = "deltaSNP", filterThreshold = 0.1, where the filter
207 threshold represents a tolerance value for filtering out relatively low confidence QTL.

208 Finally, the plotQTLstats function was used with the following parameters: var =
209 "negLog10Pval", plotThreshold = TRUE, q = 0.01, where q represents the false discovery
210 rate (FDR).

211 **3. Results and Discussion**
212

213 **3.1 Discovery of a group of spring camelina accessions that are homozygous**
214 **for functional *FLC* alleles on chromosome 20**
215

216 In Arabidopsis the *FLC* gene is required for the winter growth habit. The *FLC* gene in
217 Arabidopsis codes for a protein that has been shown to suppress the flowering promoter
218 gene, *FLOWERING LOCUS T* (*FT*) (Samach et al., 2000), and is itself suppressed (methylated)
219 by proteins that are coded by genes activated by exposure to cool temperatures (i.e. 5°C ±
220 1°C) (Sheldon et al., 2000). Polymorphisms in orthologous *FLC* alleles can differentiate the
221 spring versus winter growth habit. In the winter camelina accession Joelle (Ames 33292),
222 *FLC* is expressed on chromosomes 8 and 20, but not on chromosome 13, while its only
223 expressed on chromosome 8 in the spring camelina accession 'CO46' (Anderson et al.,
224 2018). The subfunctional allele of this gene in spring type camelina has a frameshift
225 mutation in exon 5 on chromosome 20 at bp position 4,195,043 (Anderson et al., 2018). As
226 we have looked at these alleles more broadly among a greater number of accessions within
227 the camelina germplasm phenotyped for growth habit and resequenced by Li et al., (2020)
228 (with nearly identical phenotypes reported here in Supplementary Table 1), we have found
229 that there are at least 20 exceptional cases in which a spring accession is homozygous for
230 the winter *FLC* allele on chromosome 20, two cases in which a spring accession is
231 homozygous and three cases in which a spring accession is heterozygous for the winter *FLC*
232 allele on chromosome 8, one case in which a winter accession (PI650143) is homozygous
233 for the spring *FLC* allele on chromosome 8, and one case in which a winter accession

234 (CN113692) is homozygous for the spring allele on chromosome 20 (Table 2). Winter
235 accession CN113692 was previously characterized as having a strong winter allele of *FLC* on
236 chromosome 8 and a spring allele on chromosome 20 (Chaudhary et al., 2023). Ten
237 additional camelina accessions in Table 2 were previously characterized for gene
238 expression by qRT-PCR of *FLC* on chromosomes 8 and 20 and are noted as such therein
239 (Chao et al., 2019). We found that two of these ten accessions had the spring growth habit
240 with a winter allele of *FLC* on chromosome 20, and Chao et al., (2019) found that transcript
241 abundance at this locus was relatively low for them, which explains the spring growth habit
242 of these accessions and provides a clue about the underlying genetic basis of this
243 phenomenon. Thus, when genotyping for growth habit when these exceptional spring type
244 accessions with a winter *FLC* 20 allele are parents of a given population, the qRT-PCR
245 method developed by Chao et al., (2019) is recommended. For the cases that involved a
246 heterozygous genotype where homozygosity is expected in this self-pollinating species,
247 possible explanations are that either there was a rare case of outcrossing or resequencing
248 sample contamination.

249 **3.2 Bulked segregant analysis for flowering time**
250

251 The objective of this experiment was to identify the primary chromosome specific
252 gene(s) controlling flowering time differences between the facultative winter accession of
253 camelina PI 650163-1 and a winter accession Ames 33292 (Joelle). The QTLseqr R-package
254 analysis identified just one major peak of QTL LOD score values (-$\log_{10}$ of variant p-values)
255 in the bulked segregant analysis experiment (Figure 1). The variant (SNP) with the
256 maximum LOD score (5.6) was at position 4.7 Mb on chromosome 20, was in an intergenic
257 region and thus was unlikely to influence the flowering time trait. However, it was likely in
258 linkage with the frameshift inducing INDEL in *FLC* on chromosome 20 (Anderson et al.,

259 2018), which had a LOD score of 5.4 and is at position 4.2 MB. This result supports our
260 claim that *FLC* on chromosome 20 is the primary regulator of flowering time, rather than
261 *FLC* on chromosome 8 (for which PI 650163-1 has a subfunctional allele) or any other
262 regulatory genes related to flowering time for PI 650163-1. There may be a few smaller
263 signals of SNP effects on chromosomes 4, 10, and 19, but none were comparable to the
264 signal on chromosome 20. Though the fully functional allele of *FLC* on chromosome 8 is very
265 conserved and highly expressed among winter accessions and the subfunctional allele is
266 conserved among spring accessions (Anderson et al., 2018), it appears that it plays a
267 smaller role in controlling flowering time on the growth habit level in camelina, despite the
268 fact that another report showed it has a significant effect on flowering in spring camelina
269 backgrounds (Li et al., 2020).
270

271 **3.3 Cytotype of 14 camelina accessions in the USDA-National Plant Germplasm**
272 **System.**
273

274 Previously ten accessions of *Camelina sativa* within the USDA-NPGS had been
275 cytotyped as 2n=38 and one accessions as 2n=26 by a chromosome squash and counting
276 approach (Hotton et al., 2020). This approach elucidated approximate numbers of
277 chromosomes for these *Camelina spp.* accessions, but could not positively identify the exact
278 *Camelina* species for them. We had made crosses among some of these stated 2n=38 and
279 2n=40 accessions and found that the progeny were fertile in subsequent generations (data
280 not shown). For this reason, we included eight of these accessions in a lane for whole
281 genome sequenced and then cytotyped bioinformatically compared to control accessions
282 consisting of 2n=14, 2n=26, and 2n=40 as has been described previously (Chaudhardy et al.,
283 2020). Seven of these eight we found to have the 2n=40 cytotype by the bioinformatics

284 approach, which explained why the crosses we had made among some of these accessions
285 were fertile (Figure 2, Supplementary Table 2). PI650167 was cytotyped as 2n=38 by
286 Hotton et al., and we confirmed this cytotype bioinformatically. With supporting data from
287 Chaudhary et al., 2020 and Brock et al., 2022, we were able to identify this accession as
288 *Camelina microcarpa* along with accession IHAR50003 from the IHAR germslasm system in
289 Poland, which lists it as *Camelina sativa*. Also with this supporting data, we were able to
290 identify PI650152 as *Camelina rumelica* (2n=26) along with accession CN113657, which is
291 listed as *Camelina sativa* in Plant Gene Resources of Canada.

292 **3.4 KASP genotyping approach distinguished between functional and**
293 **subfunctional alleles of *FLC* on chromosome 20**
294

295 Currently the only genotyping approach to distinguish functional from
296 subfunctional alleles of the *FLC* gene on chromosome 20 is qRT-PCR, which was developed
297 from camelina accessions we shared through an intermediary (Chao et al., 2019). Many
298 plant breeding programs make use of DNA-based Kompetitive Allele Specific PCR (KASP)
299 genotyping, so we aimed to optimize this genotyping approach for this marker. KASP
300 primers were designed for the SNV found in exon 5 of FLC on chromosome 20. While
301 individual primers in this this primer set contained many off target binding sites, these sites
302 were not adjacent to each other and thus would not allow for off target amplification to a
303 degree that could confound results. This KASP primer set was effective in distinguishing
304 winter from spring type camelina. (Figure 3 with specific camelina accessions used in
305 analysis in Supplementary Table 3).

306 **3.5 Discovery of facultative winter camelina accession PI650163**
307

308 The majority of the accessions within our camelina germplasm had a spring-type
309 growth habit (Supplementary Table 1). Based upon our fall 2015 bolting notes, we

310 conducted a follow-up greenhouse study where we planted each accession/line (20) that
311 was scored as slow bolting into pots in triplicate without a vernalizing cold treatment at any
312 point. Fourteen of these accessions did not bolt within the time frame of the experiment. PI
313 650163 repeated the slow bolting phenotype observed in fall 2015, four bolted at a
314 moderate rate, and one bolted very quickly (Figure 4). The latter may have been subjected
315 to a more stressful micro-environment in fall 2015 resulting in it being mislabeled as slow
316 bolting. By spring of 2016, none of the spring type camelina plants had survived the
317 Minnesota winter.

318 **3.6 Discovery of winter-hardy, early flowering and maturing camelina**
319 **accession PI 650163-1**
320

321 In the fall of 2017, PI650163 was planted in Rosemount, MN with the hypothesis
322 that it may be able to survive a Minnesota winter due to its facultative (slow bolting) winter
323 type growth habit. On May 18th, 2018, a single plant from this plot (PI 650163-1) was
324 selected since it was the first plant flowering in this plot as well as in the entire field. PI
325 650163-1 flowered nine and 17 days before Ames 33292 ('Joelle') in 2019 and 2020 in St.
326 Paul, MN, respectively (Table 3). Physiological maturity was also scored in St. Paul, MN in
327 2020, using the BBCH scale developed by Martinelli and Galasso (2011). PI 650163-1
328 reached 50% maturity nine days before Ames 33292 (Table 3).

329 In canola, an example of the genetic basis of the facultative winter type is increased
330 expression of a repressor of FLC - the *VENALIZATION INSENSITIVE 3* (*VIN3*) gene, which is
331 typically upregulated after exposure to temperatures near 4°C. In facultative winter type
332 canola, *VIN3* has increased expression even at a temperature of 20°C, far above what
333 typically induces its expression of this gene, and thus baseline *FLC* expression is relatively
334 lower in facultative winter type canola than in winter type (Huang et al., 2021). *Camelina*
335 *sativa* accession PI 650163 and PI 650163-1 did not have any obvious variants in *VIN3*, but

one or more other genes in this pathway may have an allele influencing this phenotype. However, since these facultative winter accessions already have a subfunctional allele of *FLC* on chromosome 20, this floral suppression pathway may be less likely to influence this phenotype. Subfunctional genes in the floral promoter pathway such as *FLOWERING LOCUS T* (*FT*) and *CONSTANS (CO)* could also be influencing this phenotype, but again, these accessions did not have any obvious variants in *FT* or in *CO*.

4 Conclusion

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

The genetic relatedness among camelina accessions showed strong clustering by within the winter and facultative winter types. Possible approaches to leverage the relatedness data could be to use it for selecting spring type parents for introducing more diversity into winter camelina breeding programs as well as to select unique parents to breed for increased seed size. Unique accessions such as PI 650143 can also be further characterized.

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

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

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

2.5 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) accessions at the University of Minnesota. We would like to thank Dr. Johann Vollmann for allowing us to study the camelina lines he developed in his breeding program at the Universität für Bodenkultur Wien (BOKU) in Vienna, Austria. We would also like to thank our undergraduate student employees, Claire Biel and Henry Parks, for their assistance with lab and field work.

Tables and Figures

Table 1. Primer design for Kompetitive Allele Specific PCR (KASP) genotyping for winter (allele X) versus spring (allele Y) growth habit in *Camelina sativa* based upon the the previously identified single basepair INDEL in exon 5 of the *FLOWERING LOCUS C* gene on chromosome 20 (Anderson et al., 2018). The bold, italic DNA sequence at the beginning of the X and Y KASP primers represent the sequences that bind to the FAM and HEX fluorescent dyes, respectively.

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

Table 2. 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)	NCBI Run	Growth Habit	FLC8 Winter Allele	FLC20 Winter Allele
PI 650163	CS229UMN	UMN Resequenced	Facultative winter	het	
PI 650163-1	CS229-1UMN	UMN Resequenced	Facultative winter		
PI 650163	CS229	SRR9993039	Facultative winter		
PI 597833	CS201	SRR9987775	Spring		
PI 652885*	CS235	SRR10028112	Spring		
PI 652886	CS236	SRR10028114	Spring		
CN 113755	CS110UMN	UMN Resequenced	Spring		x
CAM 180	CS110	SRR9914681	Spring		x
CAM 61	CS030	SRR9722050	Spring		
PI 650144	CS210	SRR10873755	Spring		
PI 311735*	CS199	SRR10873743	Spring		x
CAM 223	CS142	SRR9924268	Spring		
CAM 108	CS063	SRR9722689	Spring		x
IHAR 247026	CS211_IHAR_UMN	UMN Resequenced	Spring		
PI 650145	CS211	SRR9990156	Spring		x
PI 650146	CS212	SRR9990305	Spring		
CAM 63	CS031	SRR9722051	Spring		
PI 650140	CS206	SRR9989930	Spring		
PI 650147	CS213	SRR9990304	Spring		
PI 650152	CS218	SRR10027314	Spring		
PI 650153	CS219	SRR9990527	Spring	x	
PI 633192	CS202	SRR9988371	Spring		x
PI 633193	CS203	SRR9988367	Spring		x
PI 650161	CS227	SRR9992927	Spring		x
PI 650154	CS220	SRR10027318	Spring		
PI 650156	CS222	SRR9990625	Spring		
PI 650157B	CS223	SRR9990812	Spring		
PI 650159	CS225	SRR9992827	Spring		
PI 650160	CS226	SRR9992839	Spring		
PI 650162	CS228	SRR9993037	Spring		
PI 650164*	CS230	SRR10027313	Spring		x
PI 650165	CS231	SRR10001240	Spring		
PI 650166*	CS232	SRR10027259	Spring		
CAM 74 or 123	CS039 or 071	SRR9722741	Spring		
Ames 31231*	CS191	SRR9983193	Spring		
Ames 31232	CS192	SRR9983388	Spring		x
PI 650148	CS214	SRR9990306	Spring		
PI 650149	CS215	SRR9990308	Spring		x
PI 633194	CS204	SRR9988728	Spring		x

CAM 171	CS105	SRR9905074	Spring		x
CAM 67	CS034	SRR9722216	Spring		
PI 650150*	CS216	SRR9990309	Spring		
PI650167B	CS233	SRR10027319	Spring		
CN 113753	CS107UMN	UMN Resequenced	Spring		x
CAM 173	CS107	SRR10873422	Spring		x
Vavilov 407	UMN	UMN Resequenced	Spring	x	
CAM 7 or 57	CS003 or 027	SRR9721948	Spring		
CAM 40	CS020	SRR9721882	Spring		
PI 304268	CS195	SRR9986031	Spring		
PI 304269*	CS196	SRR9987777	Spring		
PI 650141*	CS207	SRR9989959	Spring		
?	?	SRR10873371	Spring		
CAM 25	CS007	SRR9721935	Spring		x
CAM 99	CS059	SRR9722541	Spring		
CAM 84	CS047	SRR9722430	Spring		
CAM 90	CS053	SRR9722466	Spring		
MT102	CS251	SRR9721197	Spring		
CAM 165	CS099	SRR9733108	Spring		
CAM 196	CS118	SRR9915684	Spring		
CAM 257	CS172	SRR9974986	Spring		
CAM 151	CS085	SRR9727008	Spring		
CAM 157	CS091	SRR9730513	Spring		x
MT 5	CS001	SRR9721195	Spring		x
PI 650151	CS217	SRR9990325	Spring		
CAM 68	CS035	SRR9722218	Spring		
CAM 35	CS015	SRR9721670	Spring		x
PI 258366	CS193	SRR9983359	Spring		x
CAM 110	CS064	SRR9722690	Spring		
PI 258367	CS194	SRR9986045	Spring		
?	?	SRR10873332	Spring	x	
CAM 221	CS141	SRR10873735	Spring	het	
CAM 174	CS108	SRR9914678	Spring		x
CAM 120	CS070	SRR9722738	Spring		x
CAM 78	CS041	SRR9722315	Spring	het	het
PI 650143	CS209	SRR9990157	Winter		x
PI 650155	CS221	SRR9990588	Winter	x	x
PI 650158*	CS224UMN	UMN Resequenced	Winter	x	x
PI 650168*	CS234UMN	UMN Resequenced	Winter	x	x
PI 650168	CS234	SRR10028113	Winter	x	x
PI 311736	CS200	SRR10001246	Winter	x	x
IHAR 502792	CS200-IHAR-UMN	UMN Resequenced	Winter	x	x
CAM 30	CS010	SRR9721666	Winter	no call	x
CAM 172	CS106	SRR10873423	Winter	x	x
CN 113691	CS040UMN	UMN Resequenced	Winter	x	x
CN 113692**	CS041UMN	UMN Resequenced	Winter	x	
CAM 705	CS040	SRR9722307	Winter	x	x
Ames 33292*	UMN	UMN Resequenced	Winter	x	x

*Previously evaluated for *FLC* 8, 20 winter and spring allele transcripts by Chao et al., 2019

**Previously evaluated for *FLC* 8, 20 winter and spring alleles by Chaudhary et al., 2023

Table 3. 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

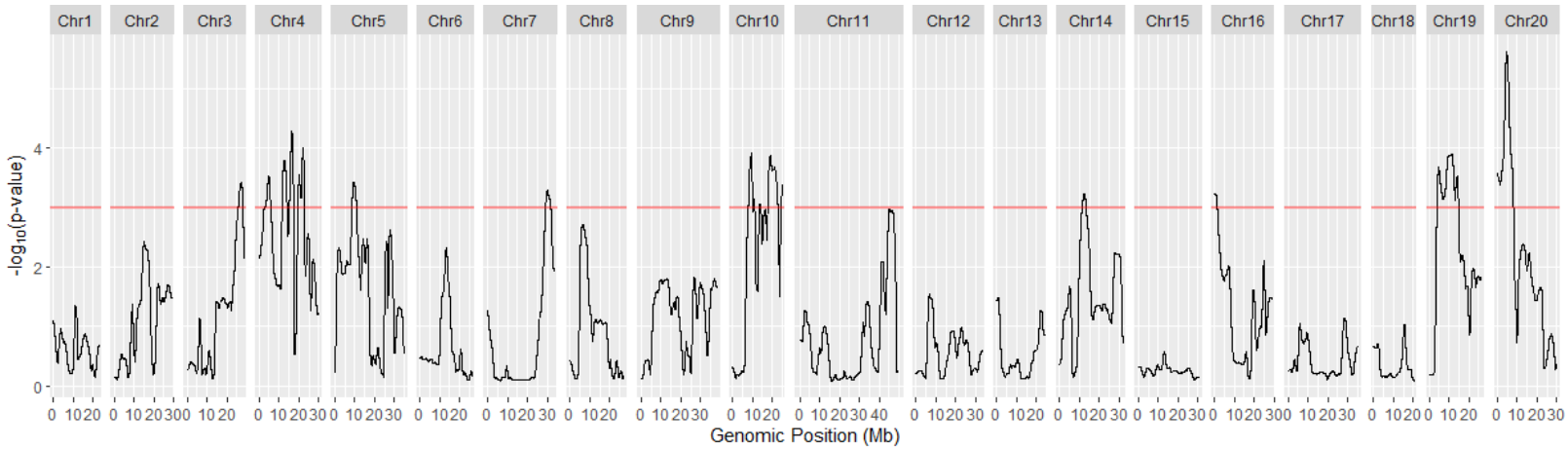

Figure 1. QTL mapping results obtained with QTLseqr and developed using bulked segregant analysis on an F_2 population of a cross between Ames 33292 (Joelle) and PI650163-1. One region on chromosome 20 in *Camelina sativa* L. Crantz is a primary regulator of flowering time. The *FLOWERING LOCUS C* gene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

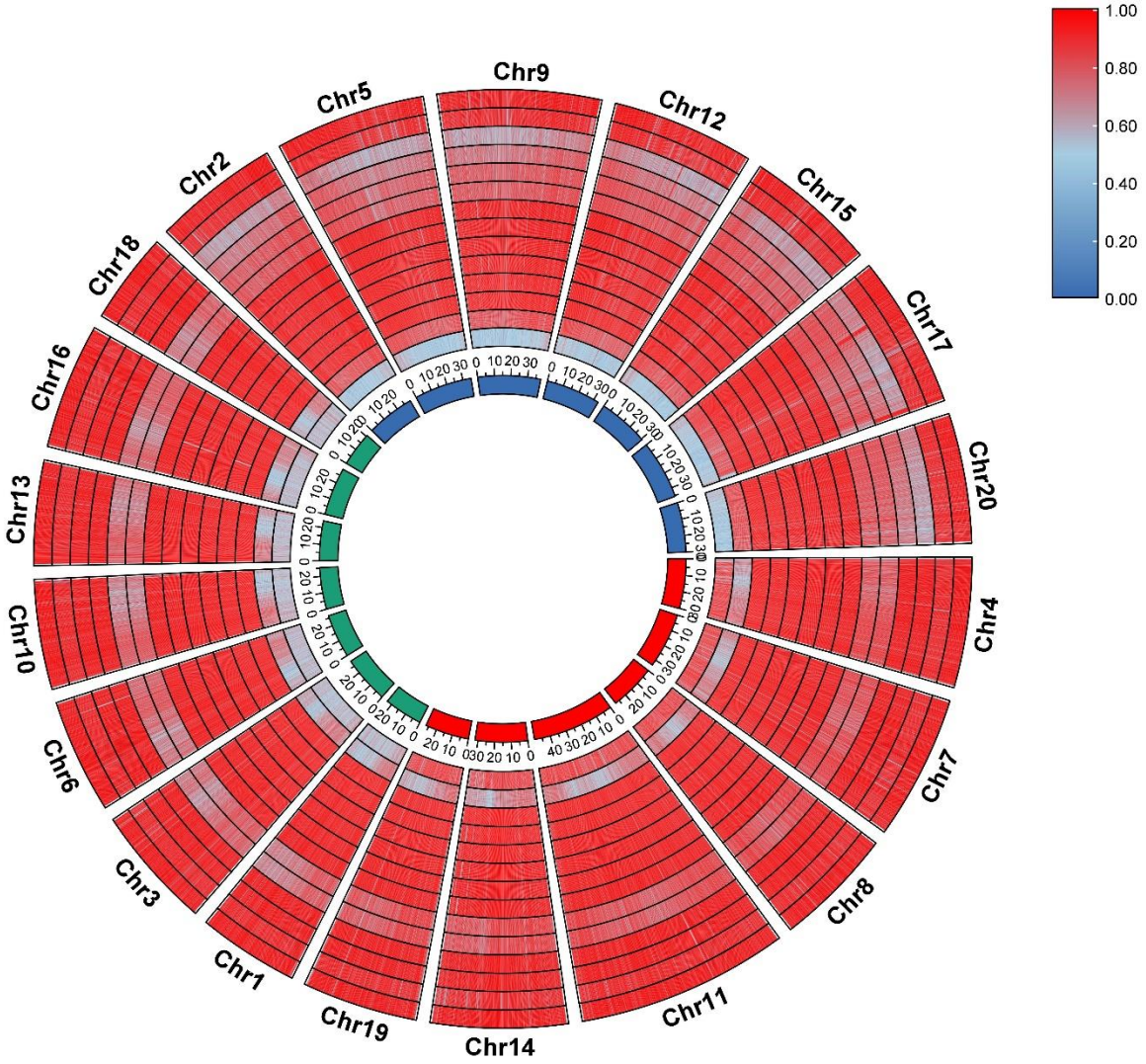

Figure 2. Circos plot heat map of the percentage of bases with non-zero coverage in 100kb windows from whole genome sequenced paired end reads aligned to the *Camelina sativa* reference genome. From the center to outside of the Circos plot: the three subgenomes of *Camelina sativa* – subgenome one in red, two in green, and three in blue. *Camelina spp*. accessions PI650135* (diploid *Camelina neglecta*), PI650133* (diploid *Camelina hispida*), and USDA-NPGS stated *Camelina sativa* accessions PI258367**, PI650141**, PI650143**, PI650144**, PI650146**, PI650164**, PI650152*** (*Camelina rumelica* 2n=26), CN113657 (*Camelina rumelica* 2n=26), PI650167**** (*Camelina microcarpa* 2n=38), IHAR50003 (*Camelina microcarpa* 2n=38), PI650168**, PI650163-1. *Previously cytotyped as diploid *C. spp.* by DNA sequencing by Chaudhary et al., 2020. **Previously cytotyped as hexaploid *Camelina* (2n=38) by chromosome squash by Hotton et al., 2020, but found to be hexaploid *Camelina sativa* (2n=40) by DNA sequencing as shown here. ***Previously cytotyped as tetraploid *Camelina* (2n=26) by chromosome squash by Hotton et al., 2020, and found to be tetraploid *Camelina rumelica* (2n=26) by DNA sequencing as shown here. ****Previously cytotyped as hexaploid *Camelina* (2n=38) by chromosome squash by Hotton et al., 2020, and found to be hexaploid *Camelina microcarpa* (2n=38) by DNA sequencing as shown here.

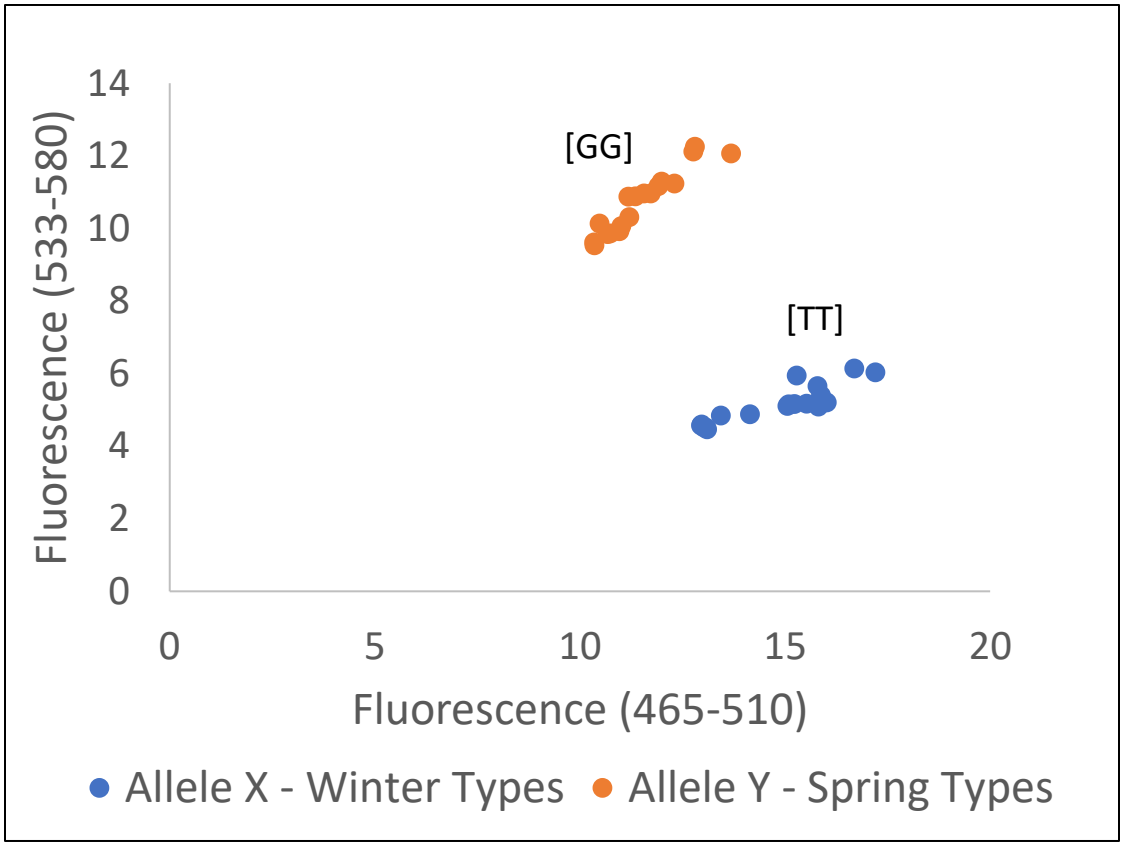

Figure 3. Kompetitive Allele-Specific PCR (KASP) genotyping results that differentiate winter from spring growth habit accessions of *Camelina sativa* using the previously identified single basepair INDEL in exon 5 of the *FLOWERING LOCUS C* gene on chromosome 20 (Anderson et al., 2018). Specific accessions used for this analysis are listed in Supplementary Table 2.

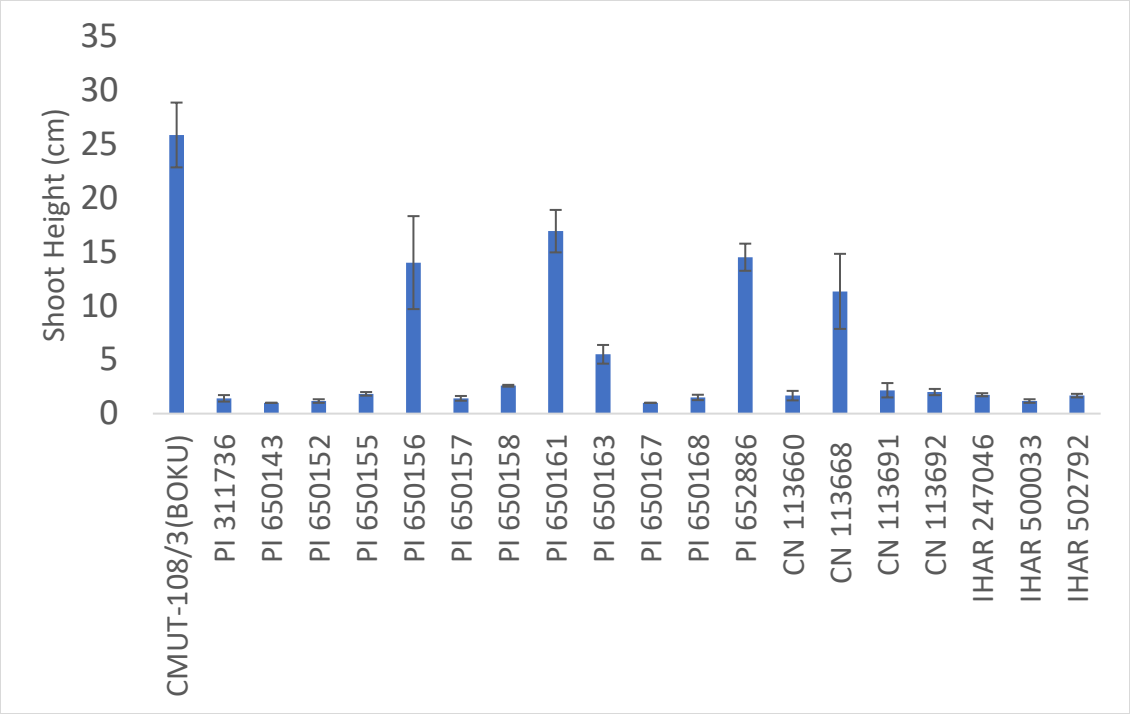

Figure 4. Camelina shoot height observed in a greenhouse 34 days after planting on February, 19th 2016 with no vernalization. Experimental design was a randomized complete block and n=3. The greenhouse temperature was maintained near 20 °C and was supplemented with halogen lighting to ensure 16h light was available each day.

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. https://doi.org/10.1002/pld3.60

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

Brock, J. R., Mandáková, T., McKain, M., Lysak, M. A., & Olsen, K. M. (2022). Chloroplast phylogenomics in Camelina (Brassicaceae) reveals multiple origins of polyploid species and the maternal lineage of C. sativa. *Horticulture Research*, *9*, uhab050. https://doi.org/10.1093/hr/uhab050

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. https://doi.org/10.1016/j.indcrop.2018.12.017

Chaudhary, R., Koh, C.S., Kagale, S., Tang, L., Wu, S.W., Lv, Z., Mason, A.S., Sharpe, A.G., Diederichsen, A. and Parkin, I.A., 2020. Assessing diversity in the Camelina genus provides insights into the genome structure of Camelina sativa. *G3: Genes, Genomes, Genetics*, *10*(4), 1297-1308. https://doi.org/10.1534/g3.119.400957

Chaudhary, R., Higgins, E.E., Eynck, C., Sharpe, A. and Parkin, I., 2023. Mapping QTL for vernalization requirement identified adaptive divergence of the candidate gene Flowering Locus C in polyploid Camelina sativa. *bioRxiv*, 2023-05. https://doi.org/10.1101/2023.05.23.541983

Chen H, Wang T, He X, Cai X, Lin R, Liang J, Wu J, King G, Wang X. 2022. [WWW Document] http://brassicadb.cn/#. Brassicaceae Database (BRAD). Nucleic Acids Res. 50 (*D1*). D1432–D1441. (accessed 3.10.2022).

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. https://doi.org/10.2134/agronj2010.0485

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. https://doi.org/10.3390/plants9050642

Huang, C.H., Sun, R.R., Hu, Y., Zeng, L.P., Zhang, N., Cai, L.M., Zhang, Q., Koch, M.A., Al-Shehbaz, I., Edger, P.P., Pires, J.C., Tan, D.Y., Zhong, Y., Ma, H., 2016. Resolution of Brassicaceae Phylogeny Using Nuclear Genes Uncovers Nested Radiations and Supports Convergent Morphological Evolution. Mol. Biol. Evol. 33, 394–412. https://doi.org/10.1093/molbev/msv226

Huang, L.Y., Min, Y., Schiessl, S., Xiong, X.H., Jan, H.U., He, X., Qian, W., Guan, C.Y., Snowdon, R.J., Hua, W., Guan, M., Qian, L.W., 2021. Integrative analysis of GWAS and transcriptome to reveal novel loci regulation flowering time in semi-winter rapeseed. PLANT Sci. 310. https://doi.org/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

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

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

Martinelli, T., Galasso, I., 2011. Phenological growth stages of Camelina sativa according to the extended BBCH scale. Ann. Appl. Biol. 158, 87–94. https://doi.org/10.1111/j.1744-7348.2010.00444.x

Ott, M.A., Eberle, C.A., Thom, M.D., Archer, D.W., Forcella, F., Gesch, R.W., Wyse, D.L., 2019. Economics and Agronomics of Relay-Cropping Pennycress and Camelina with Soybean in Minnesota. Agron. J. 111, 1281–1292. https://doi.org/10.2134/agronj2018.04.0277

Samach, A., Onouchi, H., Gold, S.E., Ditta, G.S., Schwarz-Sommer, Z., Yanofsky, M.F., Coupland, G., 2000. Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. Science (80). 288, 1613–1616. https://doi.org/10.1126/science.288.5471.1613

Sheldon, C.C., Rouse, D.T., Finnegan, E.J., Peacock, W.J., Dennis, E.S., 2000. The molecular basis of vernalization: The central role of FLOWERING LOCUS C (FLC). Proc. Natl. Acad. Sci. U. S. A. 97, 3753–3758. https://doi.org/10.1073/pnas.060023597

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

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

Wittenberg, A., Anderson, J. V, Berti, M.T., 2020. Crop growth and productivity of winter camelina in response to sowing date in the northwestern Corn Belt of the USA. Ind. Crops Prod. 158, 113036. https://doi.org/10.1016/j.indcrop.2020.113036

Zohary, D., Hopf, M., Weiss, E., 2012. Domestication of Plants in the Old World: The Origin and Spread of

Domesticated Plants in Southwest Asia, Europe, and the Mediterranean Basin. OUP Oxford.