



A novel deletion in *FLOWERING LOCUS T* modulates flowering time in winter oilseed rape

Paul Vollrath¹ · Harmeet S. Chawla¹ · Sarah V. Schiessl¹ · Iulian Gabur¹ · HueyTyng Lee¹ · Rod J. Snowdon¹ · Christian Obermeier¹

Received: 11 August 2020 / Accepted: 6 January 2021 / Published online: 20 January 2021
© The Author(s) 2021

Abstract

Key message A novel structural variant was discovered in the *FLOWERING LOCUS T* orthologue *BnaFT.A02* by long-read sequencing. Nested association mapping in an elite winter oilseed rape population revealed that this 288 bp deletion associates with early flowering, putatively by modification of binding-sites for important flowering regulation genes.

Abstract Perfect timing of flowering is crucial for optimal pollination and high seed yield. Extensive previous studies of flowering behavior in *Brassica napus* (canola, rapeseed) identified mutations in key flowering regulators which differentiate winter, semi-winter and spring ecotypes. However, because these are generally fixed in locally adapted genotypes, they have only limited relevance for fine adjustment of flowering time in elite cultivar gene pools. In crosses between ecotypes, the ecotype-specific major-effect mutations mask minor-effect loci of interest for breeding. Here, we investigated flowering time in a multiparental mapping population derived from seven elite winter oilseed rape cultivars which are fixed for major-effect mutations separating winter-type rapeseed from other ecotypes. Association mapping revealed eight genomic regions on chromosomes A02, C02 and C03 associating with fine modulation of flowering time. Long-read genomic resequencing of the seven parental lines identified seven structural variants coinciding with candidate genes for flowering time within chromosome regions associated with flowering time. Segregation patterns for these variants in the elite multiparental population and a diversity set of winter types using locus-specific assays revealed significant associations with flowering time for three deletions on chromosome A02. One of these was a previously undescribed 288 bp deletion within the second intron of *FLOWERING LOCUS T* on chromosome A02, emphasizing the advantage of long-read sequencing for detection of structural variants in this size range. Detailed analysis revealed the impact of this specific deletion on flowering-time modulation under extreme environments and varying day lengths in elite, winter-type oilseed rape.

Introduction

Structural genome variation (SV) is widespread in crop genomes and frequently associated with important traits (Gabur et al. 2019). Recent improvements in genome sequencing enable detection of SV larger than 30 base pairs

(Chawla et al. 2020). Different types of SV, including deletions, insertions, duplications, translocation and inversions (Feuk et al. 2006), can have a substantial effect on gene expression and accordingly on the phenotypes of plants (Alonge et al. 2020; Saxena et al. 2014; Wang et al. 2016). Examples include determination of different agronomic traits such as flowering time (Díaz et al. 2012; Schiessl et al. 2017b; Song et al. 2020), disease resistance (Gabur et al. 2018; Hurgobin et al. 2018), heading date (Nishida et al. 2013), seed weight (Song et al. 2020) and biotic stress response (McHale et al. 2012). Due to their cumulative size, more nucleotides tend to be affected by SV than by SNP variants on a genome-wide scale (De Coster et al. 2019).

Brassica napus ($2n = 4x = 38$, AACC) is an allotetraploid interspecific hybrid between *Brassica rapa* ($2n = 2x = 20$, AA) and *Brassica oleracea* ($2n = 2x = 18$, CC) (Chalhoub

Communicated by Maria Laura Federico.

Supplementary information The online version contains supplementary material available at (<https://doi.org/10.1007/s00122-021-03768-4>).

✉ Christian Obermeier
Christian.obermeier@agr.uni-giessen.de

¹ Department of Plant Breeding, Justus Liebig University, Giessen, Germany

et al. 2014; Parkin et al. 1995; U 1935). Due to the high similarity between the A and the C subgenomes, homoeologous exchanges occur very frequently (Chalhoub et al. 2014; Hurgobin et al. 2018; Stein et al. 2017; Szadkowski et al. 2010). Many studies to date have focused on SV in resynthesized oilseed rape, which is known to be affected by numerous and large variation. SV has been visualized using genetic mapping (Song et al. 1995; Gabur et al. 2020), molecular cytogenetics (Xiong et al. 2011), optical mapping (Gabur et al. 2020) and genome-wide short-read sequencing (Chalhoub et al. 2014; Samans et al. 2017), but all of these methods are only capable of accurately detecting large-scale SV (> 100,000 bp; Samans et al. 2017). However, using long-read sequencing, we recently demonstrated that small-scale SV events (30–10,000 bp) are also unexpectedly widespread in natural *B. napus* (Chawla et al. 2020), with considerable implications for gene function and diversity. Sequence-capture experiments demonstrated a major impact of SV on flowering-time regulatory pathway genes (Schiessl et al. 2017a, b), which play a key role in crop adaptation and yield performance. This study aims to capture the relevance of genome-wide SV in elite European winter oilseed rape and investigate the specific impact of SV on flowering time in a narrow breeding gene pool.

Facing climate change and the trend toward increasingly warmer winters in Europe, in conjunction with an earlier onset of vegetative growth after winter, flowering time is a crucial factor for plant development, seed formation and ultimately yield production. However, flowering time is determined by many genes, which complicates its examination (Blümel et al. 2015). Previous studies on flowering time in *B. napus* focused primarily on identification of major genetic factors responsible for differentiation of winter, semi-winter and spring morphotypes (Long et al. 2007; Raman et al. 2019; Schiessl et al. 2017b, 2019; Song et al. 2020; Wu et al. 2019; Xu et al. 2016). Due to the strong effects of major mutations in key flowering-time regulators, which are jointly responsible for ecogeographical differentiation, minor effects of allelic differentiation tend to be obscured in such studies. For breeding, however, where fine-tuning of flowering within ecotypes is critical, dissection of minor-effect loci in elite breeding gene pools can be more relevant than major-effect loci which disturb local adaptation. Therefore, in this study we deliberately carried out an analysis of flowering-time variation in elite winter cultivars, aiming to find novel genomic regions for fine adjustment of flowering modulation. We first investigated SV events associated with flowering-time variation in a double haploid (DH) multiparental association mapping population derived from crossing six inbred lines from elite commercial cultivars to a common parent (also a modern elite cultivar). Single-locus assays of interesting SV were further assessed for their impact on flowering time in a winter-type *B. napus* diversity

set. The results revealed associations of SV with the onset of flowering in a number of flowering-time genes within QTL regions in elite oilseed rape. The findings can help to breed for adapted oilseed rape cultivars that fulfill current and future requirements in the course of climatic changes.

Material and methods

Plant material

A multiparental population comprising 354 DH lines was used in this study (Supplementary Table S1 and Supplementary Fig. 1). The panel consists of six subpopulations derived by crossing six elite founder lines ('Adriana', 'Alpaga', 'DK Cabernet', 'Galileo', 'King 10' and the DH line 'JN') to the common elite parent 'Lorenz'. DH families from the crosses King 10 × Lorenz, Adriana × Lorenz, and JN × Lorenz were produced by NPZ Innovation GmbH (Holtsee, Germany). DH families from the crosses Lorenz × Alpaga and Lorenz × DK Cabernet were produced by Syngenta Seeds GmbH (Bad Salzuflen, Germany), while the family from the cross Lorenz × Galileo was produced by KWS SAAT SE (Einbeck, Germany). All six subpopulations comprised 60 DH lines except for the cross Lorenz × Galileo, which comprised 54 DH lines. In addition, we also investigated a set of 140 genetically diverse inbred lines of winter-type oilseed rape selected from the ERANET-ASSYST *B. napus* diversity set described in Körber et al. (2012) (Supplementary Table S2).

Field trials and phenotyping data

Field trials were conducted in an alpha lattice design with two replications at five independent locations across Germany in the 2017/2018 and 2018/2019 growing seasons. Plants were grown in plots of 6 m² (4 × 1.5 m) with a sowing density of 45 seeds per m². In 2017/18, the trials were located in Hadmersleben (Syngenta Seeds GmbH) and Nienstädt (Monsanto Agrar Deutschland GmbH), while in 2018/19 trials were carried out in Hovedissen (W. von Borries-Eckendorf GmbH & Co. KG), Leutewitz (Deutsche Saatveredelung AG) and Soerup (Bayer CropScience AG). In addition, we accessed flowering-time data from Schiessl et al. (2015) for 140 winter-type lines of the ERANET-ASSYST diversity set assessed over a total of 13 environments from 2010 to 2013, including flowering-time data from Giessen (Germany) in 2011 and 2012, Gross Gerau (Germany) in 2010, 2011 and 2012, Beibei (China) and Temuco (Chile) in 2011 and 2012, and Rauischholzhausen (Germany) in 2010, 2011, 2012 and 2013. We further assessed the diversity set for flowering time in a randomized complete block design with a plot size of 12.5 m² (10 × 1.25 m) in one replicate in

Rauischholzhausen (Germany) in 2018. Start of flowering was defined as BBCH 61 (10% of flowers on main raceme open) (Lancashire et al. 1991) and is reported in all trials as the number of days after sowing.

SNP genotyping and linkage analysis

The entire population including the seven parents was genotyped using the *Brassica* 60 k Illumina Infinium™ SNP array. By allowing zero mismatches and gaps, a stringent alignment using BLASTN (Altschul et al. 1990) was conducted and we were able to anchor 27,832 SNP uniquely to a single position of the publicly available *B. napus* genome assembly Darmor-*bzh* v4.1 (Chalhoub et al. 2014). Failed and nonspecific SNP were excluded from further analysis. Using the software Haploview, we defined linkage disequilibrium (LD) blocks based on their confidence intervals according to Gabriel et al. (2002).

Genome-wide association studies (GWAS)

The R package GenABEL (Aulchenko et al. 2007) was used to conduct a genome-wide association study (GWAS). Markers with more than 10% missing data or a minor allele frequency lower than 5% were excluded, as were genotypes with more than 10% missing data. The kinship and principal component analysis were included in the model to adjust for population stratification. We applied a false discovery rate (FDR) of ≤ 0.1 (Benjamini and Hochberg 1995) to call a marker–trait association significant. In addition, to reduce the type II error rate, a threshold of $-\log_{10}(p \text{ value}) \geq 3.0$ for marker–trait associations was defined to consider a marker as a putative SNP. Phenotypic variance (R^2) was estimated using the formula $R^2 = \frac{\chi}{n-2+\chi}$. As shown by Gabur et al. (2018), calling of single nucleotide absence polymorphisms (SNaP) by segregation patterns can provide additional markers and reveal previously undetected QTL in *B. napus*. Using this procedure, we called SNaP markers for loci that failed the threshold of < 10% failed calls but showed the expected 1:1 segregation for presence–absence polymorphisms in the respective subfamilies.

Whole genome sequencing and variant calling

Long-read sequencing on the MinION device from Oxford Nanopore Technologies (ONT) (Oxford, UK) was used to sequence genomic DNA from each parental line to a genome coverage of at least 20×. For this purpose, high molecular weight (HMW) DNA isolation was performed using a protocol modified from Mayjonade et al. (2016), as described in Chawla et al. (2020). In order to achieve a longer average read length, we conducted a size selection step using the

Circulomics Short Read Eliminator Kit (<https://www.circulomics.com/sre>) prior to the sequencing library preparation. The kit consists of a buffer that separates small DNA fragments by centrifugation. The size-selected HMW-DNA was subsequently used to prepare the sequencing library. We utilized the SQK-LSK109 sequencing kit provided by Oxford Nanopore Technologies using their recommended protocol. Library preparation consisted of the following two steps: end repair and adapter ligation, each followed by a cleaning step with magnetic beads. Finally, a library of 500–1000 ng of HMW-DNA was transferred into the MinION flow cell for sequencing. MinION raw signal data were processed using the base-caller Guppy version 3.2.1 (https://community.nanoporetech.com/downloads/guppy/release_notes). Read quality was evaluated with the tool MinIONQC (Lanfear et al. 2019), and raw reads with a minimum Q score of 7 were aligned to the *B. napus* reference genome Darmor-*bzh* v4.1 (Chalhoub et al. 2014) using the NGMLR long-read mapper version 0.2.7 (Sedlazeck et al. 2018). Subsequently, the alignment file in SAM format was converted to a sorted BAM file using SAMtools version 1.9 (Li et al. 2009) and genome-wide SV was called using Sniffles version 1.0.10 (Sedlazeck et al. 2018). Settings and subsequent filtering of SV followed Chawla et al. (2020) with a focus only on deletions and insertions. Bedtools version 2.29.2 was used to detect SV overlapping with genes (Quinlan et al. 2010). Using the Integrative Genomics Viewer version 2.6.1, we produced coverage plots and illustrated SV and genes (Robinson et al. 2011). The segregation ratio of SV was tested conducting a χ^2 test. The online database JASPAR was used to screen for transcription factor binding sites (<http://jaspar.genereg.net/>) (Khan et al. 2018).

SV validation

To validate interesting SV events detected by the variant calling, we used a standard PCR approach using primers flanking the deletions/insertions. Primer design was performed using Primer3Plus (Untergasser et al. 2007) and BLAT (Kent 2002) via the *Brassica napus* Genome Browser (<http://www.genoscope.cns.fr/brassicapopus>). All primer sequences can be found in Supplementary Table S3. Sanger sequencing of PCR products was conducted to confirm unique primer binding. Besides PCR, we confirmed the 288 bp deletion in *FLOWERING LOCUS T* on chromosome A02 (*BnaFTA02*; *BnaA02g12130D*) using 100 bp single-end Illumina short read sequencing data from 140 inbred winter-type lines of the diversity set (Schiesl et al. 2017b). Short reads from genotypes carrying the deletion fail to map accurately to *BnaFTA02* in the Darmor-*bzh* version v4.1 reference genome due to the deleted 288 bp and 92 unresolved nucleotides (Ns) within the reference. Therefore, we modified the reference by adding a corrected version of

BnaFT.A02 using long read data from cultivar Express 617 (Lee et al. 2020), once with and once without the 288 bp sequence. Alignment, removal of duplicates, sorting, indexing and normalized mean coverage (NMC) calling for the 288 bp position were performed according to Schiessl et al. (2017c). An NMC value below 0.5 indicated association with a confirmed deletion.

Results

GWAS for identification of flowering-time QTL

A total of 13,746 high-quality polymorphic SNP markers were identified in the mapping population. Additionally,

578 SNaP markers were called by analyzing the segregation patterns of raw array data in the subpopulations, as described in Gabur et al. (2018). Using all 14,324 markers, eight genomic regions associated with flowering time in winter oilseed rape were detected in at least two of five field environments (Fig. 1 and Supplementary Table S4). Four of the detected QTL were located on chromosome A02, in LD blocks with sizes of 352 kb, 498 kb, 451 kb and 383 kb. Three QTL were identified on chromosome C02 in LD blocks of 331 kb, 860 kb and 360 kb, respectively, while a 914 kb QTL interval was detected at the distal end of chromosome C03. Broad sense heritability (Falconer and Mackay 1996) for the onset of flowering was high ($h^2 = 0.83$).

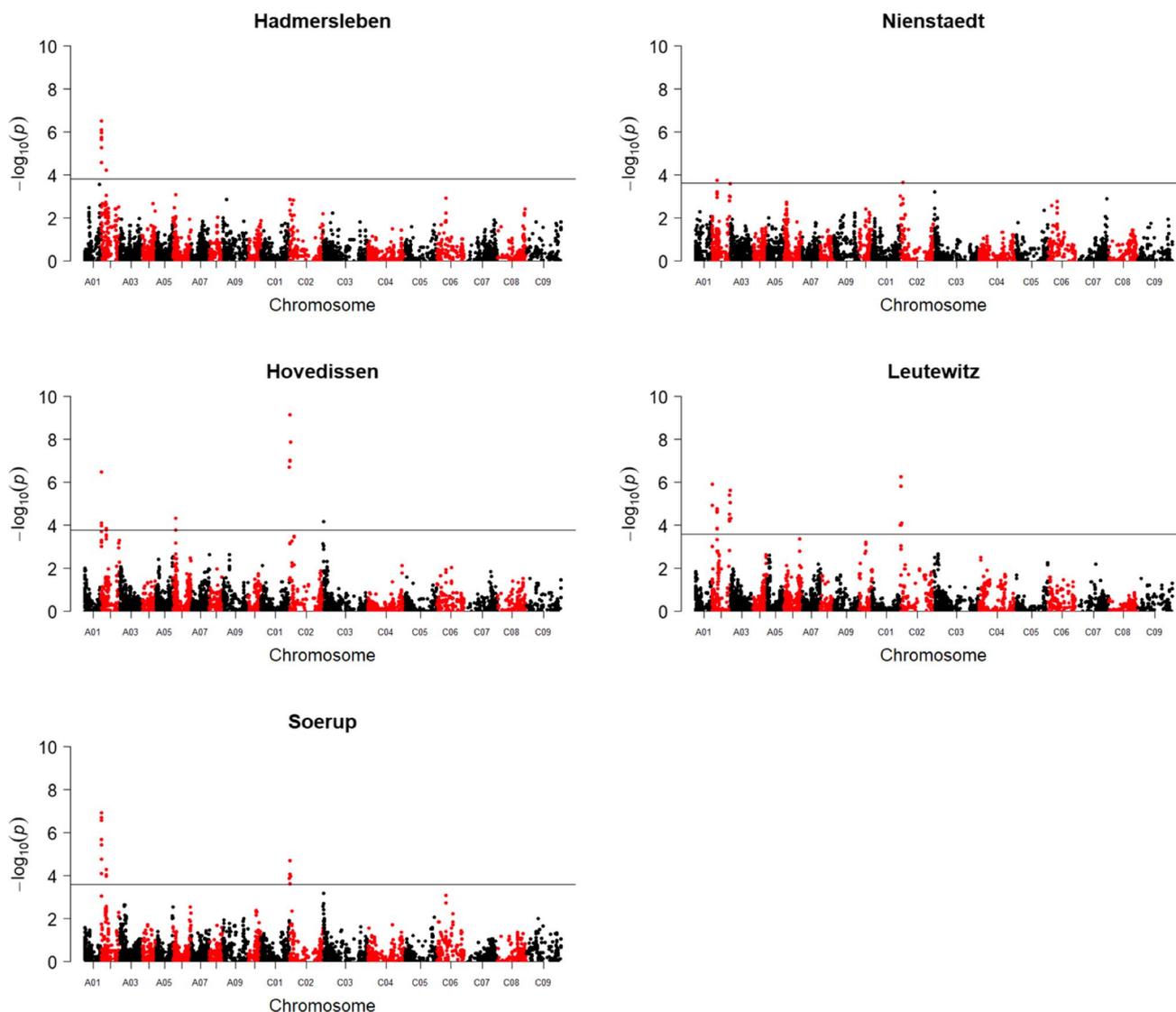


Fig. 1 Manhattan plots showing marker–trait associations for flowering time (days after sowing) in the multiparental population in five independent environments across Germany ($n = 352$). The solid line represents the threshold for significant SNP markers ($FDR \leq 0.1$)

Genome-wide structural variant detection in seven elite parental lines using long-read sequencing

A total of 250 Gb of raw sequencing data was generated from 29 MinION flowcells, with an average yield of 35.7 Gb per genotype. A genome coverage of 23× to 43× was achieved for the seven parents of the mapping population. The first 20 flowcell runs delivered average N50 values of 21,967 bp. After using the Circulomics Short Read Eliminator Kit for removing short reads from the library in the following nine sequencing runs, the N50 increased to an average of 30,526 bp, so that the final N50 of the cumulative data from all flowcells ranged from 20,418 to 29,260 bp for all seven genotypes. A summary of sequencing results including read numbers, read lengths and the quality of the reads can be found in Supplementary Table S5. Mapping rates using the Darmor-*bzh* v4.1 reference genome ranged from 63 to 72% for the seven parental lines (Supplementary Table S6). Variant calling revealed a total of 50,762 unique SV across the seven parental lines compared to Darmor-*bzh*, of which 13,374 were polymorphic across the seven parental lines. Conversely, 37,388 SV calls were consistent for all seven parental lines and therefore monomorphic in the population, but different to the Darmor-*bzh* reference; these calls likely include assembly errors of Darmor-*bzh*. Of the polymorphic SV events we detected, 6,990 were deletions and 6,384 were insertions, with sizes ranging from 30 bp to 26 kb. The majority (47.5%) had a size between 100 and 1,000 bp (Fig. 2). Median sizes for deletions and insertions were 429 bp and 263 bp, respectively. All detected polymorphic deletions and insertions were included in subsequent analyses.

Genes affected by SV within flowering-time QTL

To explore the potential impact of the genome-wide SV on protein coding genes, we firstly compared the positions of the detected SV with the gene regions for 101,040 genes annotated in the Darmor-*bzh* reference genome (Chalhoub et al. 2014). In total, 4,840 (4.8%) of all annotated genes

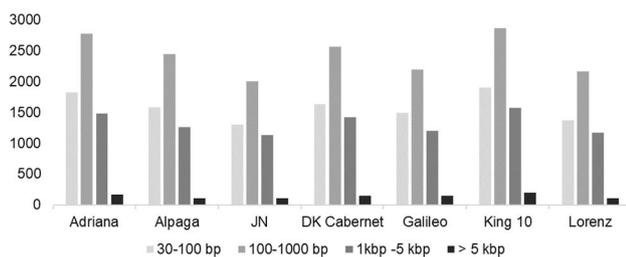


Fig. 2 Total numbers of polymorphic SV detected in seven parental lines by comparison to the *B. napus* reference Darmor-*bzh* v4.1 grouped in four different size classes

in the genome coincided with polymorphic SV events in the parental lines of the mapping population. A similar frequency of genes coinciding with SV events was observed within the eight flowering-time QTL (6.7%, 52 of 771). To investigate whether flowering time might be affected by intragenic SV, we analyzed the genes within QTL regions in more detail. In total, we found 771 annotated genes contained within the eight QTL regions, spanning 1.684 Mb on chromosome A02, 1.551 Mb on chromosome C02 and 914 kb on chromosome C03. By filtering gene ontology (GO) terms of the *B. napus* gene annotations on these 771 genes for terms that can be linked to flowering time, we identified 71 genes annotated to be involved in flowering-time regulation, among them copies of the vernalization gene *FLOWERING LOCUS C (FLC)* and the circadian rhythm gene *LATE ELONGATED HYPOCOTYL (LHY)* (Table 1 and Supplementary Table S7). Six out of these 71 genes exhibit SV. Of these six genes, five were detected in the QTL regions on chromosome A02 and one in the QTL region on chromosome C02. The affected gene on chromosome C02 is not homoeologous to any of the five flowering-related genes on chromosome A02, and the QTL regions on A02 and C02 do not represent homoeologous regions based on comparison of harbored genes as defined by Chalhoub et al. (2014). By analyzing regions up to 10 kb upstream of the start codons of the flowering-time genes, we found a large 1,361 bp deletion in the parental line JN in the promoter region of *BnaFT.A02 (BnaA02g12130D)* (Table 1). In three of the seven parental lines, the same gene also displayed a 288 bp deletion in the second intron.

Association of intragenic deletions with flowering time in the multiparental population

To evaluate trait associations with deletions in flowering-time genes, we developed specific PCR assays for all seven SV events in the six QTL-linked flowering-time candidate genes and used these to analyze the segregation of size polymorphisms identified in the parental lines in relevant subpopulations (Fig. 3, Supplementary Table S3). The seven PCR assays were first validated in the seven parental lines of the multiparental population. In 47 of these 49 PCR assays, the alleles called by ONT data from the parental lines were reconfirmed, whereas a 38 bp insertion, designated SV7, could not be reconfirmed in parental lines King 10 and Galileo and was excluded from further analyses. The PCR assays were then applied to the respective subpopulations which showed the corresponding polymorphism and tested for the expected 1:1 segregation. In most cases, the observed segregation in the subfamily did not differ significantly from the expected 1:1 segregation ratio (Table 2). However, three SV did not fit the expected segregation in 1 (from 3), 2 (from 5) and 2 (from 5) polymorphic subpopulations, respectively,

Table 1 Flowering-time candidate genes and coinciding SV identified between 7 elite parental lines within QTL for flowering time analyzed in a *B. napus* multiparental elite population ($n = 354$)

QTL name	Size of LD block in kb	Start–end position	Number of genes/flowering-time genes in Darmor- <i>bzh</i>	Number of genes/flowering-time genes with structural variants*	Gene IDs	Blast2GO annotation of gene	Name: size and type of SV (relative to Darmor- <i>bzh</i>)	PCR assay name: product size polymorphism (see Fig. 3)
A02_1	352	12,297–364,239	92/10 (10.9%)	14 / 2 (14.2%)	<i>BnaA02g00160D</i> ; <i>BnaA02g00910D</i>	Pollen development; regulation of flower development	SV1: 373 bp deletion; SV2: 64 bp deletion	P133: 1130 bp/757 bp; P255: 419 bp/362 bp
A02_2	498	6,311,854–6,810,319	59/8 (13.6%)	6/1 (16.7%)	Promoter of <i>BnaFT.A02</i> ; <i>BnaA02g12130D</i> (<i>BnaFT.A02</i>)	-; photoperiodism, flowering, positive regulation of flower development	SV3: 1,361 bp deletion; SV4: 288 bp deletion	P311: 1924 bp/563 bp; P416: 797 bp/519 bp
A02_3	451	22,569,685–23,021,163	73/3 (4.1%)	2/0 (0%)	-	-	-	-
A02_4	383	24,112,932–24,495,807	72/11 (15.3%)	5 / 2 (40%)	<i>BnaA02g33650D</i> ; <i>BnaA02g33940D</i>	Circadian rhythm; entrainment of circadian clock	SV5: 1,313 bp deletion; SV6: 35 bp deletion	P513: 819 bp/no product; P632: 667 bp/632 bp
C02_1	331	512,169–842,796	70/0 (0%)	2/0 (0%)	-	-	-	-
C02_2	860	1,028,042–1,888,045	153/11 (7.2%)	2/1 (50%)	<i>BnaC02g03640D</i>	Regulation of flower development	SV7: 38 bp insertion	P732: 262 bp/300 bp
C02_3	360	3,168,965–3,528,564	62/5 (8.1%)	10/0 (0%)	-	-	-	-
C03_1	914	2,434–915,986	190/23 (12.1%)	11/0 (0%)	-	-	-	-
Sum	4,149	-	771/71 (9.2%)	52/6 (11.5%)	-	-	-	-

Flowering-time genes were defined according to Schiessl et al. (2015) as genes associated by Blast2GO with at least one of the following gene ontology terms: flower, vernalization, photoperiod, circadian, floral, vegetative to reproductive, vegetative phase change, pollen carpel, sepal, petal

Fig. 3 Size polymorphism within six flowering-time genes from three QTL on A02 and one QTL on C02 detected by PCR in seven parents of the *B. napus* multiparental population with the common parent Lorenz. **a** SV1: 373 bp deletion, **b** SV2: 64 bp deletion, **c** SV3: 1,361 bp deletion, **d** SV4: 288 bp deletion, **e** SV5: 1,313 bp deletion, **f** SV6: 35 bp deletion, **g** SV7: 38 bp insertion (also see Tables 1 and 2 for more details)

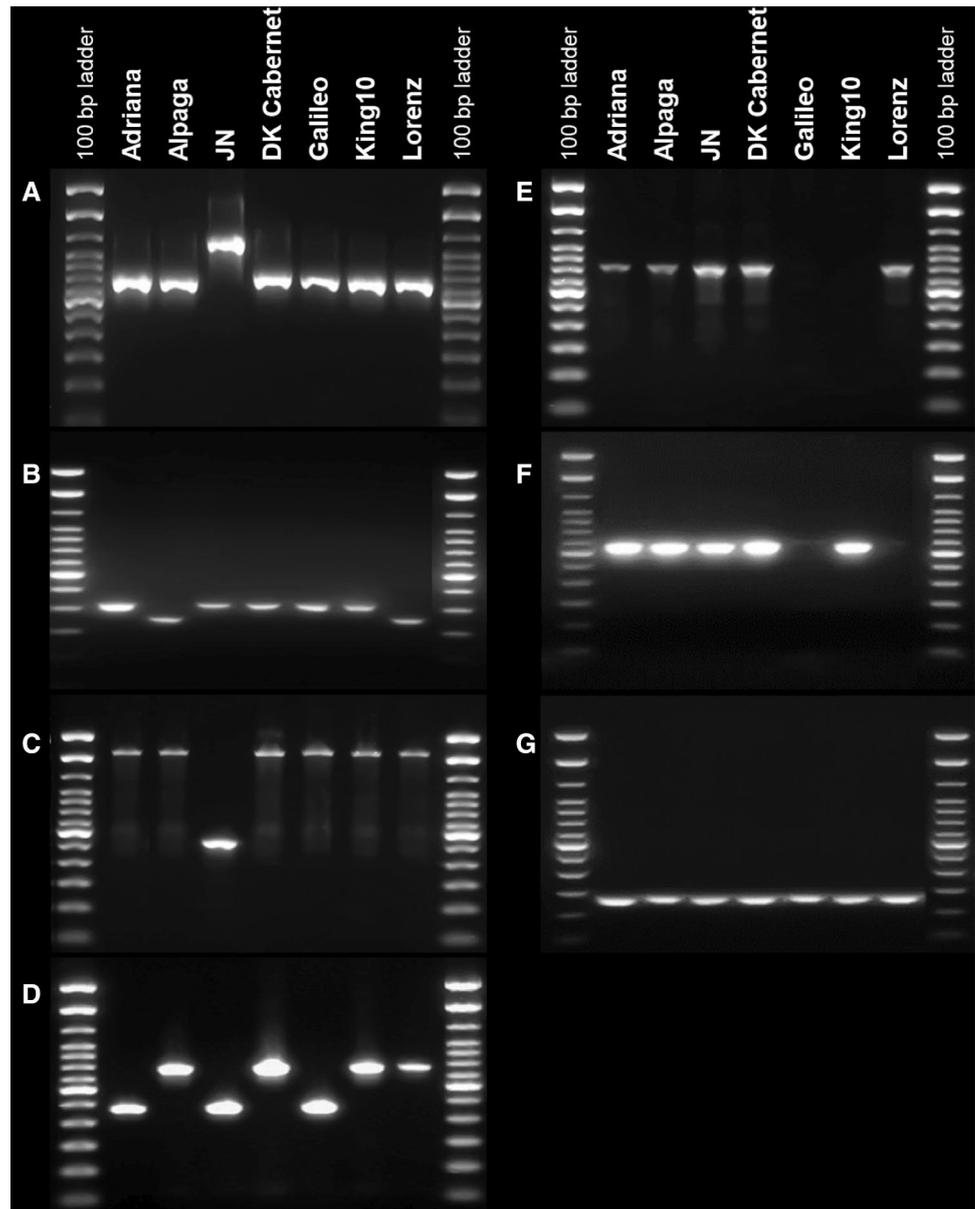


Table 2 Segregation ratios of six SV events coinciding with five flowering-time candidate genes within the investigated QTL regions of the multiparental population ($n = 354$)

SV type	SV size [in bp]	Lorenz x						
		Adriana	Alpaga	JN	DK Cabernet	Galileo	King 10	
SV1 Deletion	373	m	m	29:31 (0.796)	m	m	m	
SV2 Deletion	64	32:28 (0.606)	m	5:54 (0.000)***	28:32 (0.606)	36:18 (0.014)*	29:31 (0.796)	
SV3 Deletion	1361	m	m	31:29 (0.796)	m	m	m	
SV4 Deletion	288	43:17 (0.001)***	m	31:29 (0.796)	m	29:25 (0.586)	m	
SV5 Deletion	1313	m	m	m	m	34:20 (0.057)	34:26 (0.302)	
SV6 Deletion	35	29:31 (0.796)	39:21 (0.020)*	42:17 (0.001)**	30:30 (1.000)	m	28:32 (0.606)	

Detailed descriptions of the SV events can be found in Table 3. Numbers in brackets are the probabilities ($P[\chi^2]$) that the observed data fit the expected 1:1 segregation pattern. Asterisks show level of significance based on χ^2 test (* p value < 0.05, ** p value < 0.01, *** p value < 0.001). m: monomorphic, not tested in this subpopulation

but did fit the expected segregation in the remaining subpopulations. To understand the segregation distortions that were observed for several SV, we sequenced the PCR products. Sanger sequencing resulted in clear chromatograms, excluding the possibility of nonspecific primer binding. Particularly for allotetraploid *B. napus*, it is well known that segregation distortion in certain genome regions can occur within the process of DH production (Pilet et al. 2001; Zhao et al. 2005), either due to linkage with loci impacting DH regeneration or due to homoeologous exchanges in the F1 parent. Linked SNP markers showed the same patterns of segregation distortion, which additionally confirms a biological cause and excludes the possibility of technical issues leading to the observed distortion. However, the 60 DH lines of the subpopulations were produced from up to five single F1 plants. Thus, it is possible that the segregation distortion could be due to de novo SV in individual F1 plants. The SV calls were included in GWAS and allelic effects were calculated (Table 3). For SV4, we detected a significant association with flowering time in the multiparental population in all five locations. Effects varied between 0.38 and 1.42 days earlier onset of flowering for genotypes carrying the deletion. Interestingly, a significant effect of the SV1 deletion on flowering time (0.84–1.62 days earlier flowering) was detected in only three out of the five locations (Hadmersleben, Hovedissen and Soerup), whereas the deletion allele of SV5 showed a significant effect (0.52–1.91 days earlier flowering) only in the other two locations (Leutewitz and Nienstaedt).

A novel 288 bp deletion in *BnaFT.A02* associates with flowering time

The 288 bp deletion within the second intron of *BnaFT.A02* (SV4) was investigated in further detail because *FT* is a well described key regulator in the flowering-time pathway (Guo et al. 2014; Helliwell et al. 2006; Schiessl et al. 2017c, b; Srikanth et al. 2011) and this specific deletion has not been described before. Validation by Sanger sequencing confirmed that three of the seven parental inbred lines

(Adriana, JN, Galileo) carry the 288 bp deletion, whereas four, including the common parent Lorenz, are identical to the reference genome in this gene region. A PCR marker spanning the deletion was used to genotype the entire multiparental population for the SV polymorphism (Supplementary Table S8). Adding this PCR marker as an additional marker to the GWAS revealed a significant association (LOD score = 4.57) explaining 5.64% (R^2) of the phenotypic variation. On average, we observed a slight, but significant earlier onset of flowering (effect of deletion = -0.88 days) for genotypes having this deletion (Fig. 4). However, at location Leutewitz a stronger effect (-1.42 days) was observed (Table 3). We screened the 288 bp sequence for transcription factor binding sites using JASPAR (Khan et al. 2018).

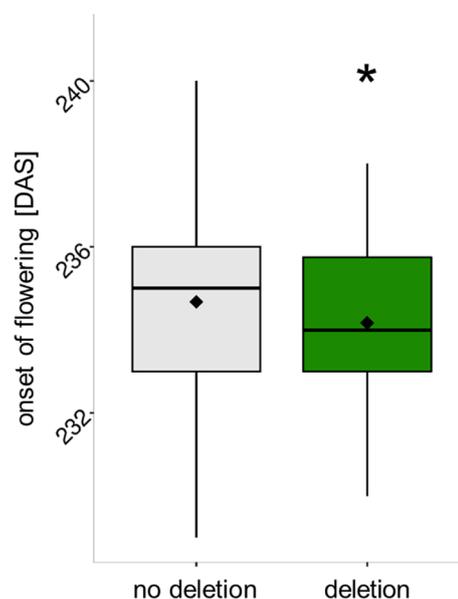


Fig. 4 Two boxplots visualize the difference for onset of flowering (days after sowing) depending on the presence/absence of the 288 bp sequence in intron 2 of the gene *BnaFT.A02*. Lines carrying the deletion show an on average earlier onset of flowering of about one day (effect of deletion = -0.88) in the *B. napus* multiparental population ($n=354$). Significant difference is highlighted with green color and asterisks (* p value < 0.05, ** p value < 0.01, *** p value < 0.001)

Table 3 Allelic effects of the six investigated SV events coinciding with five flowering-time candidate genes from the QTL in the multiparental population ($n=354$)

SV	SV type	Hadmersleben		Hovedissen		Leutewitz		Nienstaedt		Soerup	
		Effect	R^2	Effect	R^2	Effect	R^2	Effect	R^2	Effect	R^2
1	Deletion	-0.84	7.44	-1.62	4.85	ns	-	ns	-	-1.47	8.20
2	Deletion	ns	-	ns	-	ns	-	ns	-	ns	-
3	Deletion	ns	-	ns	-	ns	-	-	-	ns	-
4	Deletion	-0.38	3.63	-1.01	4.06	-1.42	3.99	-0.51	3.48	-0.80	4.84
5	Deletion	ns	-	ns	-	-1.91	6.25	-0.52	3.55	ns	-
6	Deletion	ns	-	ns	-	ns	-	ns	-	ns	-

Effect numbers describe the change in number of days to onset of flowering. R^2 : phenotypic variation (%) explained by the respective SV. ns: nonsignificant

The analysis revealed binding sites for the *CIRCADIAN CLOCK-ASSOCIATED 1 (CCA1)* and *LHY* along with several members of the REVEILLE family of transcription factors (Supplementary Table S9), suggesting that this region in the second intron of *BnaFT.A02* may have functional relevance for flowering-time modulation. None of the five other known copies of *BnaFT* in the *B. napus* genome associated to flowering-time QTL, suggesting that they are not responsible for flowering-time variation in this population (Supplementary Table S10).

Distribution of the 288 bp deletion in genetically diverse winter-type *B. napus*

The relevance of the 288 bp deletion in *BnaFT.A02* for flowering-time modulation in winter-type oilseed rape was investigated in a selection of accessions from the ERANET-ASSYST *B. napus* diversity set for which comprehensive flowering-time data were already available (Schiessl et al. 2014, 2015, 2017b). We re-analyzed Illumina short-read data produced by Schiessl et al. (2015, 2017b) from sequence capture of flowering-pathway genes, because the 288 bp deletion was not detectable in the previous analysis, where the Darmor-*bzh* v4.1 reference genome was used for alignment. However, the modified reference sequence without these 288 bp in *BnaFT.A02* allowed us to accurately map the deletion in the diversity set using the short-read sequence capture data (Supplementary Table S11). This approach confirmed the authenticity of the deletion and revealed a widespread distribution in winter-type oilseed *B. napus* cultivars. In addition, the use of short reads enabled us to confirm the borders of the deletion. Of 140 winter accessions used in this study, 34 were found to carry the deletion (NMC < 0.5), whereas 91 were similar to the reference genome (NMC > 1.5). For 15 accessions, analysis of short-read sequencing data coverage and relative copy number calculation did not reveal a clear result (NMC 0.5–1.5). We validated the approach based on short-read sequencing coverage analysis by conducting a PCR on all accessions using the same primers as described above (assay P416, Supplementary Table S3). We confirmed 32 of the 34 NMC deletion calls, whereas two accessions showed contradictory results. Out of 91 accessions without NMC deletion calls, 85 were confirmed. Two showed contradictory patterns, and four accessions were found to be heterozygous. The remaining 15 accessions include three carrying the deletion, eight without the deletion and four which were revealed to be heterozygous. This shows that relative coverage analysis based on Illumina short read data fails to accurately call a 288 bp deletion using 100 bp single-read data. For each of the 14 independent environments, a Student's *t* test was performed to calculate the phenotypic effect of the deletion assessed using the PCR assay. In this germplasm collection,

a significant association of the deletion with flowering time was found only for the locations in China and Chile in 2012, but not for the German locations. However, a trend toward early flowering associated with the deletion was observable in almost all environments (Fig. 5).

Discussion

Flowering time is an important agronomical trait that has been investigated intensively using QTL analyses and GWAS in oilseed rape. However, most studies on flowering time in *B. napus* investigated germplasm sets with broad genetic diversity including spring, semi-winter and winter-type accessions and focused on major effect QTL. Multiple flowering-time QTL with major effects and high heritabilities have been mapped, and key genes from the flowering and vernalization pathways were repeatedly implicated (Long et al. 2007; Raman et al. 2019; Song et al. 2020; Wu et al. 2019; Xu et al. 2016). In contrast, this study focused on identification of minor QTL which have not already been fixed in adapted elite breeding pools. This was achieved using a multiparental mapping population whose founders were exclusively modern, elite, European winter-type *B. napus* accessions. The aim was to identify small-effect polymorphisms relevant for modulation of flowering time in elite, adapted, winter oilseed rape breeding material. Given this fundamental difference in approach, it is not surprising that our QTL results barely overlap with previous findings. Major differences separating the three major ecotypes (spring, semi-winter and winter types) tend to mask smaller effects responsible for intra-ecotype variation. Some of the significant marker–trait associations for flowering time in the present work were stable only over a limited number of environments, suggesting that QTL effects are minor and may interact with specific environmental factors. Using 158 winter-type accessions, Schiessl et al. (2015) also revealed flowering-time QTL that are highly sensitive to local environments. In that study, none of the central flowering regulators (e.g., *Bna.FT*, *CONSTANS*, *GIGANTEA*) or central vernalization genes (e.g., *Bna.FLC*) was located within QTL regions for the onset of flowering in winter-type *B. napus*. The QTL found by Schiessl et al. (2015) also did not overlap with the QTL we found in our elite winter-type population. This indicates that fine regulation of the onset of flowering in winter-type *B. napus* is environmentally dependent and potentially controlled by multiple different genome regions with small effects.

In contrast to a previous report for disease resistance (Gabur et al. 2018, 2020), inclusion of SNaP markers (array-based presence/absence markers) in the QTL analysis did not result in detection of additional QTL associated with flowering time. This might suggest that SV may not

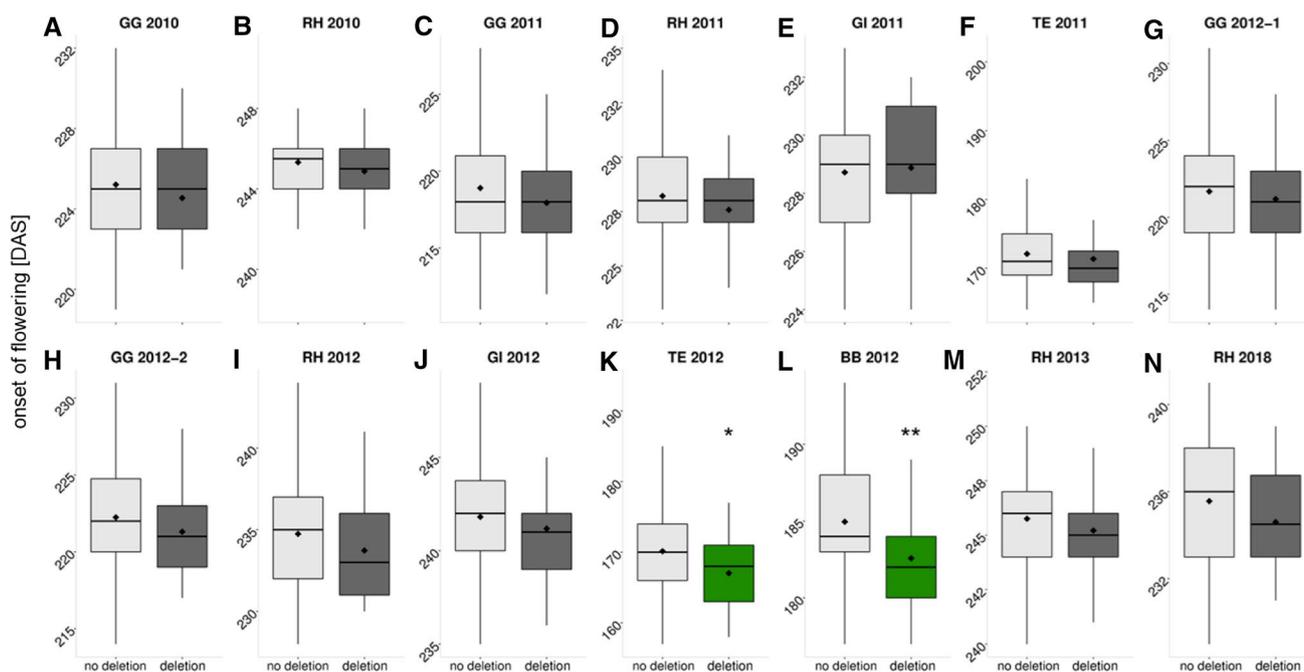


Fig. 5 The boxplots visualize the difference in the onset of flowering (days after sowing) depending on the presence/absence of the 288 bp sequence in intron 2 of the gene *BnaFTA02* in 14 environments for the ERANET-ASSYST diversity set. The environments are abbreviated as follows: Germany: Giessen (GI), Gross Gerau (GG), Raus-

chholzhausen (RH); Chile: Temuco (TE); China: Beibei, Chongqing (BB), with their respective year of harvest. Black diamonds represent the mean values. Significant differences tested with Student's *t* test are highlighted with green color and asterisks (**p* value < 0.05, ***p* value < 0.01, ****p* value < 0.001) (*n* = 140)

strongly impact flowering-time modulation in elite winter oilseed rape cultivars. However, since we showed associations between SV and flowering-time modulation, it appears more likely that SV events associated with flowering traits are in LD to neighboring SNP markers from the 60 k *Brassica* SNP array in these elite mapping parents. The number of polymorphic SV events we detected among the seven elite parental lines (13,374) is similar to the number of polymorphic SNP detected between these lines using the 60 k *Brassica* SNP array (13,746), suggesting that structural variation is a significant source of sequence polymorphism in elite cultivars with a potentially very important functional role. We found that around 5% of genes are affected by small to mid-scale SV events in each of the seven elite parents, corresponding to the frequency of intragenic SV reported by Chawla et al. (2020) for older winter oilseed rape cultivars. This underlines the importance of functional gene modification during post-polyploidization genome restructuring in *B. napus* (Chawla et al. 2020; Song et al. 2020). The high level of functionally relevant SV polymorphism in modern elite cultivars underlines the high relevance of this kind of variant for selection and breeding and the need to develop a better platform than the 60 k *Brassica* SNP array for population-wide high-throughput genotyping and identification of SV.

Most of the deletions/insertions we identified in candidate genes within flowering-time QTL in elite cultivars

reside in introns. With respect to *BnaFTA02* (*BnaA02g12130D*), one deletion lies in the second intron and another in the putative promoter region. The 1.3 kb deletion in the promoter region of *BnaFTA02* is associated with the different *B. napus* ecotypes (Chawla et al. 2020), but does not completely explain the differentiation between ecotypes (Schiessl et al. 2019). Because the populations we investigated are exclusively vernalization-dependent, this deletion did not show significant associations with flowering time. In contrast, SV1 (a 373 bp deletion in *BnaA02g00160D*), SV5 (a 1,313 bp deletion in *BnaA02g33650D*) and SV4 (a 288 bp deletion in *BnaFTA02*) showed significant associations with flowering time. Whereas a direct functional relevance of *BnaFTA02* is already well established, further functional studies are required to elucidate the potential role of the *B. napus* genes *BnaA02g00160D* and *BnaA02g33650D* in coordination of flowering time. *BnaA02g00160D* is a homolog of *A. thaliana* *SETDOMAIN GROUP 15* (*SDG15*; *AT5G09790*), a H3K27 methyltransferase which was previously described to be involved in the control of transposal activity and genome stability (Ma et al. 2018). Other methyltransferases have described roles in flowering-time regulation, like *SDG7* (Lee et al. 2015), while *SDG15* was only implicated in pollen formation so far (Raynaud et al. 2006) and is therefore considered less likely to directly

influence flowering time here. *BnaA02g33650D* is a copy of *BTB AND TAZ DOMAIN PROTEIN 1 (BT1)*, which is known to play a role in both male and female gametophyte development (Robert et al. 2009). Hence, this gene is expected to be involved in flower development, but not necessarily in flowering-time regulation, which takes place sometime prior to flower development.

Schiessl et al. (2015) did not detect the QTL harboring *BnaFT.A02* by GWAS using markers from the 60 k *Brassica* SNP array. This might be due to a strong LD decay around SV events or the lower number of accessions and the population structure of the diversity set used in this study. Moreover, the failure to detect a 288 bp deletion using a sequence capture strategy based on short single-read 100 bp Illumina sequencing (Schiessl et al. 2017b) clearly reveals the advantage of long-read sequencing to detect polymorphic small- to mid-size deletions. The 288 bp deletion in *BnaFT.A02* is associated with flowering time at all five field locations in Germany in both the 2017/2018 and 2018/2019 growing seasons. Both years were characterized by extremely warm and dry summers in Germany, in strong contrast to the environmental conditions in the years 2010–2013 during the study of Schiessl et al. (2015). This may suggest a QTL-by-environment interaction that is more active under extreme environments, which in turn might explain why day length is only crucial in the diversity set.

FT is one of the key regulators of flowering time in the well-investigated flowering-time pathway (Helliwell et al. 2006; Srikanth et al. 2011). Several studies, particularly in *Arabidopsis*, elucidated interactions of *FT* with other major genes in this pathway (Faure et al. 2007; Turck et al. 2008). For instance, the single copy of *FT* in *Arabidopsis* is suppressed by *FLC*, which is downregulated during extended periods of cold temperature to unlock the repression (Helliwell et al. 2006; Searle et al. 2006). However, the allopolyploid *B. napus* genome possesses multiple *FT* orthologs and paralogs with incompletely elucidated function. For example, *BnaFT.A02* is most likely not regulated by *Bna.FLC*. Instead, Wang et al. (2009) showed that the characteristic CArG box to which *FLC* normally binds is missing in this particular *FT* copy, and Guo et al. (2014) demonstrated that *BnaFT.A02* is expressed independently of vernalization. On the other hand, *FT* expression is activated by *CONSTANS (CO)*, whose activity is promoted by an increase in photoperiod (Kobayashi et al. 1999; Srikanth et al. 2011; Turck et al. 2008; Wigge et al. 2005). In agreement with Schiessl et al. (2015), who demonstrated the day length dependency of *BnaFT.A02*, we found that a deletion polymorphism in *BnaFT.A02* in winter-type *B. napus* only associated with flowering time during trials in Chile and China, where day lengths differed significantly from the German trial locations. This supports a vernalization-independent,

day-length-dependent activation of *BnaFT.A02* (Corbesier et al. 2007; Guo et al. 2014; Schiessl et al. 2015).

Expression of *FT* can be detected in leaves and apical meristems (Wigge et al. 2005). Corbesier et al. (2007) proposed the *FT* protein moves from the leaves through the phloem to the meristem. Changes in *FT* expression have been shown to significantly change flowering time in *Arabidopsis* (Kobayashi et al. 1999) as well as in major crops like barley (Faure et al. 2007), wheat (Yan et al. 2006), maize (Lazakis et al. 2011; Meng et al. 2011) and potato (Navarro et al. 2011). In winter-type *B. napus*, *Bna.FT* expression changes caused by allelic variants were associated with altered flowering time (Tudor et al. 2020), while high *FT* expression is known to correlate with earlier flowering (Raman et al. 2019). To investigate potential mechanisms by which the 288 bp deletion within the second intron of *BnaFT.A02* might regulate expression of this *FT* ortholog, we searched the sequence spanning the SV for putative regulatory motifs known to be involved in controlling gene expression. Besides the main ‘*FT*-interactors’ *FLC* and *CO*, several other genes are already known to interact with *FT*. For example, *SCHLAFMÜTZE* (Mathieu et al. 2009), *TEMPRANILLO 1* (Castillejo et al. 2008) and *SHORT VEGETATIVE PHASE* (Kobayashi et al. 2007) repress *BnaFT.A02* expression by binding to it. Screening for transcription factor binding sites within the deleted 288 bp sequence revealed binding sites for the well-known circadian clock genes *CCA1* and *LHY*. In addition, *LHY* was found to be located in QTL C03_1. We also noted two unannotated genes with similarity to known flowering-time genes, *GA20-oxidase (BnaC02g01710D)* and *CONSTANS-LIKE (BnaC02g06280D)*, within QTL C02_1 and C02_3, respectively. This suggests that this SV polymorphism may modulate *FT* expression by altering the binding characteristics to one or both of these genes, which are known to act together by the formation of dimers (Lu et al. 2009; Seo et al. 2012) and which have known involvement in photoperiod responses (Fujiwara et al. 2008; Li et al. 2008).

In the past, the prevalence of SV within genes has been mostly ignored in most crop species due to the lack of resolution of short-read resequencing. Recent breakthroughs in long-range sequencing technologies have led to increased application of these tools in recent years, particularly for whole-genome assembly (Michael et al. 2018; Schmidt et al. 2017). To a limited extent, long-read resequencing has also been applied to explore genome-wide SV patterns; however, the focus to date was generally on broad comparisons between small numbers of very diverse accessions (e.g., Chawla et al. 2020; Todesco et al. 2020), or strongly restructured genomes of mapping parents with synthetic background (Gabur et al. 2019). Associations of SV with flowering time, disease resistance and eco-geographical differentiation in diverse oilseed rape *B. napus* accessions (Chawla et al. 2020) revealed a general relevance of SV for

breeding. However, the extent, relevance and role of SV for important traits in modern elite cultivars have still not been addressed or documented in detail. Here, we show that SV in key regulator genes of flowering time is associated with flowering-time modulation in elite winter oilseed rape cultivars. The present study stresses the potential value of SV impacting important agronomic traits within a set of well-adapted and strongly selected varieties. Our results suggest that the utilization of markers developed from SV can add additional benefit to breeding programs. Our findings can be used to fine-adjust commercial varieties in terms of flowering time, an increasingly important adaptation trait in the face of climate change.

Acknowledgements We thank Regina Illgner for excellent technical help. We thank all German breeding companies organized in the GFPi (Gemeinschaft zur Förderung von Pflanzeninnovation e. V.) for production of the elite multiparental population and for performing field trials and providing phenotyping data.

Author Contribution statement CO and RS conceived the idea. PV, HSC and CO developed the methodology. CO and RS sourced the funding. SVS provided phenotype and reanalyzed Illumina sequencing data. PV performed the laboratory analysis. PV and CO carried out data curation. PV, HSC, SVS, IG, HTL and CO analyzed the data. PV and CO drafted the manuscript. PV, HSC, SVS, IG, HTL, RS and CO revised and gave further inputs to this manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. PV and CO received funding from the German Federal Ministry of Education and Research (BMBF) within the project PhomaDur (Grant Number BLE: 2818205115). This work was also supported by the BMBF-funded de.NBI Cloud within the German Network for Bioinformatics Infrastructure (de.NBI) (031A537B, 031A533A, 031A538A, 031A533B, 031A535A, 031A537C, 031A534A, 031A532B).

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alonge M, Wang X, Benoit M, Soyk S, Pereira L, Zhang L, Suresh H, Ramakrishnan S, Maumus F, Ciren D, Levy Y, Harel TH, Shalev-Schlosser G, Amsellem Z, Razifard H, Caicedo AL, Tieman DM, Klee H, Kirsche M, Aganezov S, Ranallo-Benavidez TR, Lemmon ZH, Kim J, Robitaille G, Kramer M, Goodwin S, McCombie WR, Hutton S, van Eck J, Gillis J, Eshed Y, Sedlazeck FJ, van der Knaap E, Schatz MC, Lippman ZB (2020) Major impacts of widespread structural variation on gene expression and crop improvement in tomato. *Cell* 182:146–161. <https://doi.org/10.1016/j.cell.2020.05.021>
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Aulchenko YS, de Koning D-J, Haley C (2007) Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. *Genetics* 177:577–585. <https://doi.org/10.1534/genetics.107.075614>
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289–300. <https://doi.org/10.2307/2346101>
- Blümel M, Dally N, Jung C (2015) Flowering time regulation in crops—what did we learn from *Arabidopsis*? *Curr Opin Biotechnol* 32:121–129. <https://doi.org/10.1016/j.copbio.2014.11.023>
- Castillejo C, Pelaz S (2008) The balance between CONSTANS and TEMPRANILLO activities determines FT expression to trigger flowering. *Curr Biol* 18:1338–1343. <https://doi.org/10.1016/j.cub.2008.07.075>
- Chalhoub B, Denoeud F, Liu S, Parkin IAP, Tang H, Wang X, Chiquet J, Belcram H, Tong C, Samans B, Corréa M, Da Silva C, Just J, Falentin C, Koh CS, Le Clainche I, Bernard M, Bento P, Noel B, Labadie K, Alberti A, Charles M, Arnaud D, Guo H, Daviaud C, Alamy S, Jabbari K, Zhao M, Edger PP, Chelaifa H, Tack D, Lassalle G, Mestiri I, Schnel N, Le Paslier M-C, Fan G, Renault V, Bayer PE, Golicz AA, Manoli S, Lee T-H, Thi VHD, Chalabi S, Hu Q, Fan C, Tollenaere R, Lu Y, Bataill C, Shen J, Sidebottom CHD, Wang X, Canaguier A, Chauveau A, Bérard A, Deniot G, Guan M, Liu Z, Sun F, Lim YP, Lyons E, Town CD, Bancroft I, Wang X, Meng J, Ma J, Pires JC, King GJ, Brunel D, Delourme R, Renard M, Aury J-M, Adams KL, Batley J, Snowdon RJ, Tost J, Edwards D, Zhou Y, Hua W, Sharpe AG, Paterson AH, Guan C, Wincker P (2014) Plant genetics. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345:950–953. <https://doi.org/10.1126/science.1253435>
- Chawla HS, Lee H, Gabur I, Vollrath P, Tamilselvan-Nattar-Amutha S, Obermeier C, Schiessl SV, Song J-M, Liu K, Guo L, Parkin IAP, Snowdon RJ (2020) Long-read sequencing reveals widespread intragenic structural variants in a recent allopolyploid crop plant. *Plant Biotechnol J*. <https://doi.org/10.1111/pbi.13456>
- Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C, Coupland G (2007) FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. *Science* 316:1030–1033. <https://doi.org/10.1126/science.1141752>
- De Coster W, De Rijk P, De Roeck A, De Pooter T, D'Hert S, Strazisar M, Kristel S, van Broeckhoven C (2019) Structural variants identified by oxford nanopore promethion sequencing of the human genome. *Genome Res* 29:1178–1187. <https://doi.org/10.1101/gr.244939.118>
- Díaz A, Zikhali M, Turner AS, Isaac P, Laurie DA (2012) Copy number variation affecting the Photoperiod-B1 and Vernalization-A1 genes is associated with altered flowering time in wheat (*Triticum aestivum*). *PLoS ONE* 7:e33234. <https://doi.org/10.1371/journal.pone.0033234>
- Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th edn. Prentice Hall, London
- Faure S, Higgins J, Turner A, Laurie DA (2007) The *FLOWERING LOCUS T*-like Gene Family in Barley (*Hordeum*

- vulgare*). Genetics 176:599–609. <https://doi.org/10.1534/genetics.106.069500>
- Feuk L, Marshall CR, Wintle RF, Scherer SW (2006) Structural variants: changing the landscape of chromosomes and design of disease studies. Hum Mol Genet 15:R57–66. <https://doi.org/10.1093/hmg/ddl057>
- Fujiwara S, Oda A, Yoshida R, Niinuma K, Miyata K, Tomozoe Y, Tajima T, Nakagawa M, Hayashi K, Coupland G, Mizoguchi T (2008) Circadian clock proteins *LHY* and *CCA1* regulate SVP protein accumulation to control flowering in *Arabidopsis*. Plant Cell 20:2960–2971. <https://doi.org/10.1105/tpc.108.061531>
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structure of haplotype blocks in the human genome. Science 296:2225–2229. <https://doi.org/10.1126/science.1069424>
- Gabur I, Chawla HS, Liu X, Kumar V, Faure S, von Tiedemann A, Jestin C, Dryzka E, Volkmann S, Breuer F, Delourme R, Snowdon R, Obermeier C (2018) Finding invisible quantitative trait loci with missing data. Plant Biotechnol J 16:2102–2112. <https://doi.org/10.1111/pbi.12942>
- Gabur I, Chawla HS, Lopisso DT, von Tiedemann A, Snowdon RJ, Obermeier C (2020) Gene presence-absence variation associates with quantitative *Verticillium longisporum* disease resistance in *Brassica napus*. Sci Rep 10:4131. <https://doi.org/10.1038/s41598-020-61228-3>
- Gabur I, Chawla HS, Snowdon RJ, Parkin IAP (2019) Connecting genome structural variation with complex traits in crop plants. Theor Appl Genet 132:733–750. <https://doi.org/10.1007/s00122-018-3233-0>
- Guo Y, Harloff HJ, Jung C, Molina CM (2014) Mutations in single *FT*- and *TFL1*-paralogs of rapeseed (*Brassica napus* L.) and their impact on flowering time and yield components. Front Plant Sci 5:282. <https://doi.org/10.3389/fpls.2014.00282>
- Helliwell CA, Wood CC, Robertson M, James Peacock W, Dennis ES (2006) The arabidopsis FLC protein interacts directly in vivo with *SOC1* and *FT* chromatin and is part of a high-molecular-weight protein complex. Plant J 46:183–192. <https://doi.org/10.1111/j.1365-3113X.2006.02686.x>
- Hurgobin B, Golicz AA, Bayer PE, Chan C-KK, Tirnaz S, Dolatabadian A, Schiessl SV, Samans B, Montenegro JD, Parkin IAP, Pires JC, Chalhoub B, King GJ, Snowdon R, Batley J, Edwards D (2018) Homoeologous exchange is a major cause of gene presence/absence variation in the amphidiploid *Brassica napus*. Plant Biotechnol J 16:1265–1274. <https://doi.org/10.1111/pbi.12867>
- Kent WJ (2002) BLAT—the BLAST-like alignment tool. Genome Res 12:656–664. <https://doi.org/10.1101/gr.229202>
- Khan A, Fornes O, Stigliani A, Gheorghie M, Castro-Mondragon JA, van der Lee R, Bessy A, Chêneby J, Kulkarni SR, Tan G, Baranasic D, Arenillas DJ, Sandelin A, Vandepoele K, Lenhard B, Ballester B, Wasserman WW, Parcy F, Mathelier A (2018) JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucleic Acids Res 46:D1284. <https://doi.org/10.1093/nar/gkx1188>
- Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T (1999) A pair of related genes with antagonistic roles in mediating flowering signals. Science 286:1960–1962. <https://doi.org/10.1126/science.286.5446.1960>
- Kobayashi Y, Weigel D (2007) Move on up, it's time for change—mobile signals controlling photoperiod-dependent flowering. Genes Dev 21:2371–2384. <https://doi.org/10.1101/gad.1589007>
- Körber N, Wittkop B, Bus A, Friedt W, Snowdon RJ, Stich B (2012) Seedling development in a *Brassica napus* diversity set and its relationship to agronomic performance. Theor Appl Genet 125:1275–1287. <https://doi.org/10.1007/s00122-012-1912-9>
- Lancashire PD, Bleiholder H, van den Boom T, Langelüddeke P, Stauss R, Weber E, Witzemberger A (1991) A uniform decimal code for growth stages of crops and weeds. Ann Appl Biol 119:561–601. <https://doi.org/10.1111/j.1744-7348.1991.tb04895.x>
- Lanfear R, Schalamun M, Kainer D, Wang W, Schwesinger B (2019) MinIONQC: fast and simple quality control for MinION sequencing data. Bioinformatics 35:523–525. <https://doi.org/10.1093/bioinformatics/bty654>
- Lazakis CM, Coneva V, Colasanti J (2011) *ZCN8* encodes a potential orthologue of *Arabidopsis FT* florigen that integrates both endogenous and photoperiod flowering signals in maize. J Exp Bot 62:4833–4842. <https://doi.org/10.1093/jxb/err129>
- Lee H, Chawla HS, Obermeier C, Dreyer F, Abbadi A, Snowdon R (2020) Chromosome-scale assembly of winter oilseed rape *Brassica napus*. Front Plant Sci 496:1–13. <https://doi.org/10.3389/fpls.2020.00496>
- Lee J, Yun J-Y, Zhao W, Shen W-H, Amasino RM (2015) *SDG7* and proper timing of vernalization. PNAS 112:2269–2274. <https://doi.org/10.1073/pnas.1423585112>
- Li D, Liu C, Shen L, Wu Y, Chen H, Robertson M, Helliwell CA, Ito T, Meyerowitz E, Yu H (2008) A repressor complex governs the integration of flowering signals in *Arabidopsis*. Dev Cell 15:110–120. <https://doi.org/10.1016/j.devcel.2008.05.002>
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Long Y, Shi J, Qiu D, Li R, Zhang C, Wang J, Hou J, Zhao J, Shi L, Park B-S, Choi SR, Lim YP, Meng J (2007) Flowering time quantitative trait loci analysis of oilseed *Brassica* in multiple environments and genomewide alignment with *Arabidopsis*. Genetics 177:2433–2444. <https://doi.org/10.1534/genetics.107.080705>
- Lu SX, Knowles SM, Andronis C, Ong MS, Tobin EM (2009) *CIRCADIAN CLOCK ASSOCIATED1* and *LATE ELONGATED HYPOCOTYL* function synergistically in the circadian clock of *Arabidopsis*. Plant Physiol 150:834–843. <https://doi.org/10.1104/pp.108.133272>
- Ma Z, Castillo-González C, Wang Z, Sun D, Hu X, Shen X, Potok ME, Zhang X (2018) Arabidopsis serrate coordinates histone methyltransferases *ATXR5/6* and RNA processing factor *RDR6* to regulate transposon expression. Dev Cell 45:769–784. <https://doi.org/10.1016/j.devcel.2018.05.023>
- Mathieu J, Yant LJ, Mürdter F, Küttner F, Schmid M (2009) Repression of flowering by the miR172 target *SMZ*. PLoS Biol 7:e1000148. <https://doi.org/10.1371/journal.pbio.1000148>
- Mayjonade B, Gouzy J, Donnadieu C, Pouilly N, Marande W, Callot C, Langlade N, Muñoz S (2016) Extraction of high-molecular-weight genomic DNA for long-read sequencing of single molecules. Biotechniques 61:203–205. <https://doi.org/10.2144/000114460>
- McHale LK, Haun WJ, Xu WW, Bhaskar PB, Anderson JE, Hyten DL, Gerhardt DJ, Jeddloh JA, Stupar RM (2012) Structural variants in the soybean genome localize to clusters of biotic stress-response genes. Plant Physiol 159:1295–1308. <https://doi.org/10.1104/pp.112.194605>
- Meng X, Muszynski MG, Danilevskaia ON (2011) The *FT*-like *ZCN8* Gene functions as a floral activator and is involved in photoperiod sensitivity in maize. Plant Cell 23:942–960. <https://doi.org/10.1105/tpc.110.081406>
- Michael TP, Jupe F, Bemm F, Motley ST, Sandoval JP, Lanz C, Loudet O, Weigel D, Ecker JR (2018) High contiguity *Arabidopsis thaliana* genome assembly with a single nanopore flow cell. Nat Commun 9:541. <https://doi.org/10.1038/s41467-018-03016-2>

- Navarro C, Abelenda JA, Cruz-Oro E, Cuéllar CA, Tamaki S, Silva J, Shimamoto K, Prat S (2011) Control of flowering and storage organ formation in potato by *FLOWERING LOCUS T*. *Nature* 478:119–122. <https://doi.org/10.1038/nature10431>
- Nishida H, Yoshida T, Kawakami K, Fujita M, Long B, Akashi Y, Laurie DA, Kato K (2013) Structural variation in the 5' upstream region of photoperiod-insensitive alleles *Ppd-A1a* and *Ppd-B1a* identified in hexaploid wheat (*Triticum aestivum* L.), and their effect on heading time. *Mol Breed* 31:27–37. <https://doi.org/10.1007/s11032-012-9765-0>
- Parkin IA, Sharpe AG, Keith DJ, Lydiate DJ (1995) Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). *Genome* 38:1122–1131. <https://doi.org/10.1139/g95-149>
- Pilet ML, Duplan G, Archipiano M, Barret P, Baron C, Horvais R, Tanguy X, Lucas MO, Renard M, Delourme R (2001) Stability of QTL for field resistance to blackleg across two genetic backgrounds in oilseed rape. *Crop Sci* 41:197–205. <https://doi.org/10.2135/cropsci2001.411197x>
- Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842. <https://doi.org/10.1093/bioinformatics/btq033>
- Raman H, Raman R, Qiu Y, Yadav AS, Sureshkumar S, Borg L, Rohan M, Wheeler D, Owen O, Menz I, Balasubramanian S (2019) GWAS hints at pleiotropic roles for *FLOWERING LOCUS T* in flowering time and yield-related traits in canola. *BMC Genom* 20:636. <https://doi.org/10.1186/s12864-019-5964-y>
- Raynaud C, Sozzani R, Glab N, Domenichini S, Perennes C, Cella R, Kondorosi E, Bergounioux C (2006) Two cell-cycle regulated SET-domain proteins interact with proliferating cell nuclear antigen (PCNA) in *Arabidopsis*. *Plant J* 47:395–407. <https://doi.org/10.1111/j.1365-313X.2006.02799.x>
- Robert HS, Quint A, Brand D, Vivian-Smith A, Offringa R (2009) BTB and TAZ domain scaffold proteins perform a crucial function in *Arabidopsis* development. *Plant J* 58:109–121. <https://doi.org/10.1111/j.1365-313X.2008.03764.x>
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP (2011) Integrative genomics viewer. *Nat Biotechnol* 29:24–26. <https://doi.org/10.1038/nbt.1754>
- Samans B, Chalhoub B, Snowdon RJ (2017) Surviving a genome collision: genomic signatures of allopolyploidization in the recent crop species *Brassica napus*. *Plant Genome* 10:1–15. <https://doi.org/10.3835/plantgenome2017.02.0013>
- Saxena RK, Edwards D, Varshney RK (2014) Structural variations in plant genomes. *Brief Funct Genom* 13:296–307. <https://doi.org/10.1093/bfpg/elu016>
- Schiessl S, Huettel B, Kuehn D, Reinhardt R, Snowdon R (2017a) Post-polyploidisation morphotype diversification associates with gene copy number variation. *Sci Rep* 7:41845. <https://doi.org/10.1038/srep41845>
- Schiessl S, Huettel B, Kuehn D, Reinhardt R, Snowdon RJ (2017b) Targeted deep sequencing of flowering regulators in *Brassica napus* reveals extensive copy number variation. *Sci Data* 4:170013. <https://doi.org/10.1038/sdata.2017.13>
- Schiessl S, Iniguez-Luy F, Qian W, Snowdon RJ (2015) Diverse regulatory factors associate with flowering time and yield responses in winter-type *Brassica napus*. *BMC Genom* 16:737. <https://doi.org/10.1186/s12864-015-1950-1>
- Schiessl S, Samans B, Hüttel B, Reinhardt R, Snowdon RJ (2014) Capturing sequence variation among flowering-time regulatory gene homologs in the allopolyploid crop species *Brassica napus*. *Front Plant Sci* 2014:1–14. <https://doi.org/10.3389/fpls.2014.00404>
- Schiessl SV, Huettel B, Kuehn D, Reinhardt R, Snowdon RJ (2017c) Flowering time gene variation in *Brassica* Species shows evolutionary principles. *Front Plant Sci* 2017:1–13. <https://doi.org/10.3389/fpls.2017.01742>
- Schiessl SV, Quezada-Martinez D, Tebartz E, Snowdon RJ, Qian L (2019) The vernalisation regulator *FLOWERING LOCUS C* is differentially expressed in biennial and annual *Brassica napus*. *Sci Rep* 9:14911. <https://doi.org/10.1038/s41598-019-51212-x>
- Schmidt MH-W, Vogel A, Denton AK, Istace B, Wormit A, van de Geest H, Bolger ME, Alseekh S, Maß J, Pfaff C, Schurr U, Chetelat R, Maumus F, Aury J-M, Koren S, Fernie AR, Zamir D, Bolger AM, Usadel B (2017) De Novo assembly of a new *Solanum pennellii* accession using nanopore sequencing. *Plant Cell* 29:2336–2348. <https://doi.org/10.1105/tpc.17.00521>
- Searle I, He Y, Turck F, Vincent C, Fornara F, Kröber S, Amasino RA, Coupland G (2006) The transcription factor *FLC* confers a flowering response to vernalization by repressing meristem competence and systemic signaling in *Arabidopsis*. *Genes Dev* 20:898–912. <https://doi.org/10.1101/gad.373506>
- Sedlazeck FJ, Rescheneder P, Smolka M, Fang H, Nattestad M, von Haeseler A, Schatz MC (2018) Accurate detection of complex structural variations using single-molecule sequencing. *Nat Methods* 15:461–468. <https://doi.org/10.1038/s41592-018-0001-7>
- Seo PJ, Park M-J, Lim M-H, Kim S-G, Lee M, Baldwin IT, Park C-M (2012) A Self-Regulatory Circuit of *CIRCADIAN CLOCK-ASSOCIATED1* underlies the circadian clock regulation of temperature responses in *Arabidopsis*. *Plant Cell* 24:2427–2442. <https://doi.org/10.1105/tpc.112.098723>
- Song J-M, Guan Z, Hu J, Guo C, Yang Z, Wang S, Liu D, Wang B, Lu S, Zhou R, Xie W-Z, Cheng Y, Zhang Y, Liu K, Yang Q-Y, Chen L-L, Guo L (2020) Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. *Nat Plants* 6:34–45. <https://doi.org/10.1038/s41477-019-0577-7>
- Song K, Lu P, Tang K, Osborn TC (1995) Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proc Natl Acad Sci U S A* 92:7719–7723. <https://doi.org/10.1073/pnas.92.17.7719>
- Srikanth A, Schmid M (2011) Regulation of flowering time: all roads lead to Rome. *Cell Mol Life Sci* 68:2013–2037. <https://doi.org/10.1007/s00018-011-0673-y>
- Stein A, Coriton O, Rousseau-Guétin M, Samans B, Schiessl SV, Obermeier C, Parkin IAP, Chèvre A-M, Snowdon RJ (2017) Mapping of homoeologous chromosome exchanges influencing quantitative trait variation in *Brassica napus*. *Plant Biotechnol J* 15:1478–1489. <https://doi.org/10.1111/pbi.12732>
- Szadkowski E, Eber F, Huteau V, Lodé M, Huneau C, Belcram H, Coriton O, Manzanares-Dauleux MJ, Delourme R, King GJ, Chalhoub B, Jenczewski E, Chèvre A-M (2010) The first meiosis of resynthesized *Brassica napus*, a genome blender. *New Phytol* 186:102–112. <https://doi.org/10.1111/j.1469-8137.2010.03182.x>
- Todesco M, Owens GL, Bercovich N, Lègaré J-S, Soudi S, Burge DO, Huang K, Ostevik KL, Drummond EBM, Imerovski I, Lande K, Pascual-Robles MA, Nanavati M, Jahani M, Cheung W, Staton SE, Muñoz S, Nielsen R, Donovan LA, Burke JM, Yeaman S, Rieseberg LH (2020) Massive haplotypes underlie ecotypic differentiation in sunflowers. *Nature*. <https://doi.org/10.1038/s41586-020-2467-6>
- Tudor EH, Jones DM, He Z, Bancroft I, Trick M, Wells R, Irwin JA, Dean C (2020) QTL-seq identifies *BnaFTA.02* and *BnaFLCA.02* as candidates for variation in vernalisation requirement and response in winter oilseed rape (*Brassica napus*). *Plant Biotechnol J*. <https://doi.org/10.1111/pbi.13421>
- Turck F, Fornara F, Coupland G (2008) Regulation and identity of florigen: *FLOWERING LOCUS T* moves center stage. *Annu Rev Plant Biol* 59:573–594. <https://doi.org/10.1146/annurev.arpla.59.032607.092755>
- U N, (1935) Genome analysis in *Brassica* with special reference to the experimental formation of *Brassica napus* and peculiar mode of fertilization. *Jap J Bot* 7:389–453

- Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM (2007) Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res* 35:W71–W74. <https://doi.org/10.1093/nar/gkm306>
- Wang J, Long Y, Wu B, Liu J, Jiang C, Shi L, Zhao J, King GJ, Meng J (2009) The evolution of *Brassica napus* *FLOWERING LOCUS T* paralogues in the context of inverted chromosomal duplication blocks. *BMC Evol Biol* 9:271. <https://doi.org/10.1186/1471-2148-9-271>
- Wang X, Wang H, Liu S, Ferjani A, Li J, Yan J, Yang X, Qin F (2016) Genetic variation in *ZmVPP1* contributes to drought tolerance in maize seedlings. *Nat Genet* 48:1233–1241. <https://doi.org/10.1038/ng.3636>
- Wigge PA, Kim MC, Jaeger KE, Busch W, Schmid M, Lohmann JU, Weigel D (2005) Integration of spatial and temporal information during floral induction in *Arabidopsis*. *Science* 309:1056–1059. <https://doi.org/10.1126/science.1115983>
- Wu D, Liang Z, Yan T, Xu Y, Xuan L, Tang J, Zhou G, Lohwasser U, Hua S, Wang H, Chen X, Wang Q, Zhu Le, Maodzeka A, Hussain N, Li Z, Li X, Shamsi IH, Jilani G, Wu L, Zheng H, Zhang G, Chalhoub B, Shen L, Yu H, Jiang L (2019) Whole-genome resequencing of a worldwide collection of rapeseed accessions reveals the genetic basis of ecotype divergence. *Mol Plant* 12:30–43. <https://doi.org/10.1016/j.molp.2018.11.007>
- Xiong Z, Gaeta RT, Pires JC (2011) Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*. *Proc Natl Acad Sci U S A* 108:7908–7913. <https://doi.org/10.1073/pnas.1014138108>
- Xu L, Hu K, Zhang Z, Guan C, Chen S, Hua W, Li J, Wen J, Yi B, Shen J, Ma C, Tu J, Fu T (2016) Genome-wide association study reveals the genetic architecture of flowering time in rapeseed (*Brassica napus* L.). *DNA Res* 23:43–52. <https://doi.org/10.1093/dnares/dsv035>
- Yan L, Fu D, Li C, Blechl A, Tranquilli G, Bonafede M, Sanchez A, Valarik M, Yasuda S, Dubcovsky J (2006) The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *Proc Natl Acad Sci U S A* 103:19581–19586. <https://doi.org/10.1073/pnas.0607142103>
- Zhao J, Becker HC, Zhang D, Zhang Y, Ecke W (2005) Oil content in a European × Chinese rapeseed population: QTL with additive and epistatic effects and their genotype-environment interactions. *Crop Sci* 45:51–59

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