Mol Breeding (2020) 40: 3
https://doi.org/10.1007/s11032-019-1087-z

Fine-mapping of a major quantitative trait locus *Qdff3-1* controlling flowering time in watermelon

Winnie Gimode · Josh Clevenger · Cecilia McGregor

Received: 9 August 2019 / Accepted: 26 November 2019 / Published online: 20 December 2019

Abstract Flowering time is crucial in watermelon (*Citrullus lanatus*) production as it determines time of fruit set. Early flowering is desirable because it enables crops to escape biotic and abiotic stresses that are intensified by long production cycles. Production of seedless watermelon is also reliant on synchronized flowering of diploid pollenizers and the triploid watermelon cultivars. Utilizing marker-assisted selection (MAS) of flowering time in watermelon breeding would potentially aid in selection for early flowering, which would shorten the production time. A major quantitative trait locus *Qdff3-1* (12–17 Mb) associated with days to female flower (R^2 = 50%) was previously identified on chromosome 3 of watermelon. In this study, we validated the *Qdff3-1* locus using QTL-seq. To determine more precisely the interval of *Qdff3-1* and the candidate gene controlling flowering time, SNP markers were identified in the region and Kompetitive Allele Specific PCR (KASP) assays were developed for high-throughput genotyping. Markers were tested for trait association on the mapping population, recombinant $F_{2:3}$ populations, and a panel of differential cultivars. In the KBS x NHM genetic background the QTL was delineated to a 1.13-Mb region, flanked by markers UGA3_14537958 and NW0248748. This region includes the *FT* and a protein phosphatase 2C (*PP2C*) gene. Genotyping the regions of interest in a panel of genetically diverse cultivars suggests that genetic control of flowering time in watermelon is dependent on the genetic background. These results lay the foundation for a greater understanding of flowering mechanisms in watermelon and improved breeding strategies for this trait.

Keywords *Citrullus lanatus* · Flowering time · *FT* · MAS · SNP · *PP2C* · QTL-seq

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11032-019-1087-z) contains supplementary material, which is available to authorized users.

W. Gimode
Institute for Plant Breeding, Genetics & Genomics, University of Georgia, 1111 Plant Sciences Bldg, Athens, GA 30602, USA

J. Clevenger
Institute for Plant Breeding, Genetics & Genomics, University of Georgia, 111 Riverbend Rd. CAGT, Athens, GA 30602, USA

C. McGregor (✉)
Department of Horticulture and Institute for Plant Breeding, Genetics & Genomics, University of Georgia, 1111 Plant Sciences Bldg, Athens, GA 30602, USA
e-mail: cmcgre1@uga.edu

Introduction

Flowering represents the transition from the vegetative to the reproductive growth phase of a plant (Levy and Dean 1998) and is essential for reproductive success (Zimmerman 1988). Flowering at the appropriate time is crucial because pollination and subsequent seed and fruit development depend on it (Jung and Müller 2009). In watermelon production, flowering is an important trait as it determines time of fruit set and therefore earliness of a cultivar. Earliness is desirable because it allows crops to escape diseases and other abiotic stresses

that intensify later in the season (Poland et al. 2009). In addition, cultivars that produce flowers early in the season can result in shorter production times. Through the use of early flowering cultivars, grower input costs are significantly reduced and their profits increased through earlier harvesting dates (Mohr 1986).

Due to consumer demand for seedless fruit in the USA, there has been an increase in the production of triploid watermelon to approximately 90% of the US watermelon market (U.S. Department of Agriculture 2017). Diploid pollenizers are required for successful production of seedless watermelon because triploid plants used to produce the seedless fruit lack sufficient viable pollen (Maynard and Elmstrom 1992). Synchronized flowering time between the pollenizers and triploid cultivars is thus essential in seedless watermelon production (McGregor and Waters 2013).

The production cycle in watermelon is determined by the time at which the first male and first female flowers are produced, and the interval between the opening of these flowers (McGregor et al. 2014; Wehner 2008). Temperature and photoperiod are major factors that determine when a plant will flower and set seed (Maheswaran et al. 2000) and in watermelon, this is promoted by high light intensity and temperature (Wehner 2008). Approximately one female flower is produced for every seven male flowers in watermelon (Wehner 2008); however, this depends on the cultivar and the environmental conditions.

Several genes have been found to control flowering in *Arabidopsis*, and pathways that control floral transition have been determined. These pathways include the photoperiodic, autonomous, vernalization, and hormonal (gibberellin) pathways (Mouradov et al. 2002; Poethig 2003; Putterill et al. 2004; Simpson and Dean 2002). In many species of flowering plants, flowering locus T (*FT*) appears to be one of the major genes that regulates flowering (Ahn et al. 2006; Kardailsky et al. 1999; Kobayashi et al. 1999; Schwartz et al. 2009) and it has been found to be part of the signaling system in cucurbits as well (Lin et al. 2007; Lu et al. 2014).

MAS for flowering time has been useful in breeding programs of several economically important crops. Quantitative trait loci (QTL) related to flowering time have been identified in various crops such as maize, wheat, soybean, and rice (Ducrocq et al. 2009; Maheswaran et al. 2000; Thomson et al. 2003; Yan et al. 2006; Yano et al. 2001). A major gene, *Hd3a*, which is an ortholog of *Arabidopsis FT*, had a functional mutation associated with days to flowering in rice (Ordonez et al. 2010). In barley and wheat, a favorable allele of the vernalization gene (*VRN3*) was found to promote flowering, while in maize, the gene *dlfl* is hypothesized to interact with an *FT* ortholog to promote flowering (Ducrocq et al. 2009). In tomato, two QTL associated with days to flowering (DTF) were detected (Jimenez-Gomez et al. 2007; Sumugat et al. 2010) and an allele of the gene *SFT*, which is an ortholog of *FT* in *Arabidopsis*, promoted flowering (Lifschitz et al. 2006). Flowering time QTL have also been mapped in Brassica crops (Bohuon et al. 1998; Ferreira et al. 1995; Osborn et al. 1997; Rae et al. 1999; Robert et al. 1998) and genes orthologous to *CONSTANS* (*CO*) in *Arabidopsis* were identified as candidate genes (Robert et al. 1998). In cucumber, *Csa1G651710*, which is an ortholog of the *FT* gene in *Arabidopsis*, was also identified as a possible candidate gene for early flowering (Lu et al. 2014). Although QTL associated with flowering time have been identified in most crops, information on specific genes and functional mutations associated with the trait remain elusive.

A major QTL associated with days to male flower (DMF (*Qdmf3-1*)) and days to female flower (DFF (*Qdff3-1*)), which contributed approximately 50% of the phenotypic variance in the Klondike Black Seeded (KBS; late flowering) × New Hampshire Midget (NHM; early flowering) recombinant inbred line (RIL) population ($n = 145$), was previously identified on chromosome 3 of watermelon (McGregor et al. 2014). This QTL had a LOD-1.5 support interval stretching from 8.2 to 12 cM (≈12–17 Mb) (Guo et al. 2013; McGregor et al. 2014) and a significant positive correlation between DMF and DFF (0.79, $P < 0.0001$) was observed in this population. One hundred seventy-two genes were predicted within this region (Guo et al. 2013), including *Cla009504*, an ortholog of the *FT* gene associated with flowering time in other species (Kojima et al. 2002; Lin et al. 2007; Lu et al. 2014), and *Cla000855*, an ortholog of the TEMPRANILLO (*TEM1*) gene in *A. thaliana* (Castillejo and Pelaz 2008). *TEM1* has been described as a transcription factor, regulated by the circadian clock which acts as a repressor of *FT* (Castillejo and Pelaz 2008).

Understanding the genetic control of flowering time in watermelon may reduce the production time by enabling selection for early flowering. It also has applications in seedless watermelon breeding as it could inform selection of the most suitable pollenizers for the triploid

cultivars. This study specifically focused on DFF, since fruit set occurs on the female flower. Due to the high correlation between DMF and DFF, however, association between these two phenotypes would be expected. In this study, we validated and refined the previously identified flowering time locus (*Qdff3-1*) using the QTL-seq method (Takagi et al. 2013), and we developed KASP markers to span the region, used $F_{2:3}$ populations to further fine-map the region, and tested the markers on a diverse cultivar panel.

Materials and methods

Plant materials

The F_6 KBS × NHM RIL mapping population used in the study was developed from a cross between KBS and NHM watermelon cultivars (Sandlin et al. 2012) with contrasting flowering time phenotypes (McGregor et al. 2014). The F_2 populations used for fine-mapping were obtained from selfing four different F_1 plants from the same cross.

DNA extraction and candidate gene sequencing

The two candidate genes identified within the previously identified *Qdff3-1* locus by Sandlin et al. (2012), i.e., *Cla009504* (*FT*) and *Cla000855* (*TEM1*), were sequenced (Eurofins Genomics LLC, Louisville, KY) in KBS and NHM. Genomic DNA was extracted from approximately 50 mg of young leaf tissue frozen in liquid nitrogen using the E.Z.N.A. Plant DNA kit (Omega Bio-Tek Inc., Norcross, GA) following the manufacturer instructions. The DNA was quantified using a Tecan NanoQuant Plate™ (Tecan, Group Ltd., Mannedorf, Switzerland) compatible with an Infinite M200pro reader (Tecan Group Ltd.). PCR amplifications were carried out in 20 µl containing 12.6 µl ddH_2O and 0.6 µl DMSO, and final concentrations of 1× Phusion HF Buffer, 0.2 mM dNTPs, 0.5 µM of forward and reverse primers (Supplementary Table S1), 0.4 U/µl Phusion Taq Polymerase (New England BioLabs Inc., Ipswich, MA), and 100 ng/µl genomic DNA. PCR conditions consisted of an initial incubation at 98 °C for 30 s, 35 cycles at 98 °C for 5 s, primer annealing temperature for 20 s, and 72 °C for 30 s, followed by 1 cycle of 72 °C for 10 min and then held at 4 °C. Prior to sequencing, amplification of PCR products were confirmed by gel electrophoresis using 3 µl of the product on a 1.5% agarose gel. The PCR products were then purified using the E.Z.N.A. Cycle Pure kit (Omega Bio-Tek Inc.). Sanger sequencing was performed for two plants of each cultivar at Eurofins Genomics LLC. Sequences were quality trimmed then aligned to the 97103 watermelon genome (Guo et al. 2013) using Geneious 11.1.4 (http://www.geneious.com).

QTL-seq

To validate the location of the *Qdff3-1* locus and identify additional SNPs for fine-mapping, QTL-seq (Takagi et al. 2013) was performed.

Phenotyping

The KBS × NHM RIL population (n = 159) previously used to map flowering time (McGregor et al. 2014) was used for this study. During summer 2016, the population, parental, and F_1 plants were transplanted at the Durham Horticulture Farm in Watkinsville, GA. Using a randomized complete block design (RCBD) with 5 blocks, one plant per RIL and 2 plants per parental cultivar/F_1 were planted per block. Plants were grown on plastic mulch with between-row spacing of 1.83 m and in-row spacing of 1.22 m. Data was collected three times a week, for days to anthesis for the first female flower (DFF).

Bulk construction and DNA extraction

RILs for each bulk were selected from the KBS × NHM population based on the average of the 2012 (McGregor et al. 2014) and 2016 flowering time data. The 2013 data described in McGregor et al. (2014) was not used due to the unusually high amount of rainfall in the period after transplanting during that year, which had an effect on flowering. Two DNA bulks, i.e., "early" (E-bulk: DFF = 21 ± 0.8) and "late" (L-bulk: DFF = 32 ± 0.9), were constructed using DNA from 15 RILs each. DNA of the selected individuals was extracted and quantified as previously described for the sequenced samples. Equal amounts of DNA from the 15 respective individuals were bulked and shipped for library construction and 151-bp paired-end whole genome Illumina sequencing on the Illumina HiSeq X (Illumina, San Diego, CA) at the Hudson Alpha Genomic services laboratory (Huntsville, AL).

Analysis of short reads and SNP identification

FastQC (Andrews, 2010) was used to analyze the quality of the short reads obtained. The 97103 watermelon genome sequence (Guo et al. 2013) was indexed using BWA, after which reads were aligned using BWA MEM (Li et al. 2009). The reads were then converted from SAM to BAM format using SAMtools (Li et al. 2009), and the converted files were sorted and indexed. Genotype likelihoods were calculated using SAMtools and SNP calling was performed using BCFtools. The SNPs were then filtered with a minimum depth of 10 using a custom python script. A SNP-index value for each SNP was obtained by dividing the number of reads matching the reference by the total number of reads for that SNP. This was calculated separately for the "early" (E) and "late" (L) bulks. The ΔSNP-index was then calculated by subtracting the SNP-indices (E-bulk and L-bulk) of the bulks (Takagi et al. 2013). A threshold for $P < 0.01$ was calculated for the population, taking into account the population size, number of individuals in each bulk, and read depth. To identify regions of the genome highly associated with flowering time, 1-Mb sliding window analysis with a 10-Kb step increment was applied.

KASP marker development and genotyping

DNA was extracted from leaf tissue of the RIL mapping population using a salt extraction method (King et al. 2014) with the following modifications: The tissue sample plates stored at − 80 °C were placed in liquid nitrogen prior to using a TissueLyser II (QIAGEN, Hilden, Germany) for sample grinding. Five hundred microliters of the SDS/NaCl extraction buffer (40% (v/v) 5 M NaCl and 60% (v/v) extraction buffer (200 mM Tris/HCl pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% SDS)) was used and plates were centrifuged for 10 min at 3600 rpm.

KASP primers (Supplementary Table S2) for SNPs identified through candidate gene sequencing and QTL-seq were designed and optimized using Primer3Plus (Untergasser et al. 2007). The KASP primers spanning the QTL region were designed for high-throughput genotyping of the populations and cultivars and were named with the prefix "UGA3_" followed by the physical position of the SNP on the chromosome in Mb. KASP primers for SNPs in the genetic map described by Sandlin et al. (2012) (NW0247945:12,804,860 Mb and NW0248748: 15,664,017 Mb) were not renamed. KASP assays were carried out in 4-μl volumes containing 1.94 μl of 2 × low rox KASP master mix (LGC Genomics LLC, Teddington, UK), 0.06 μl primer mix with a final primer concentration of 0.81 μM, and 2 μl of 50–100 ng/μl genomic DNA. The PCR conditions used for the KASP assays consisted of an initial incubation at 95 °C for 15 min, 10 cycles of touchdown PCR with 20 s at 95 °C, 25 s of primer annealing temperature + 9 °C with 1 °C decrease each cycle, and 15 s of 72 °C, followed by 35 cycles of 10 s at 95 °C, 1 min at primer annealing temp, and 15 s at 72 °C, then held at 4 °C. KASP fluorescent end-point readings were measured using an Infinite M200Pro plate reader (Tecan Group Ltd.) and genotype calls were made using KlusterCaller (LGC Genomics LLC). Flapjack version 1.18.06.29 (Milne et al. 2010) was used for haplotype visualization. To determine the marker/trait association, a one-way ANOVA followed by a Tukey-Kramer HSD test was used. R^2 values of each marker were tested using the Kruskal-Wallace test ($P = 0.05$) to determine the association of genotypes with flowering time in the mapping population and cultivar panel.

Fine-mapping the QTL interval for *Qdff3-1*

To determine more precisely the interval of *Qdff3-1* and the candidate gene(s) controlling the trait, recombinants were identified. From the F_6 KBS × NHM RIL population, one line (RIL-190) carrying a recombination in one of the major recombination points was selected. Three hundred seventy-six seedlings from four KBS × NHM F_2 populations were genotyped using KASP markers from different haplotype blocks identified in the F_6 RIL population. Six recombinants (RIL-190 and five F_2 lines) were selected to contain heterozygous or homozygous genotypes at three different recombination points. The recombinant plants were selfed and the resulting seeds were sown in the greenhouse and the seedlings transplanted 5 weeks later at the UGA Durham Horticulture farm for flowering time evaluation in summer 2018. The populations, consisting of approximately 70 plants each and 7 plants each of parents and F_1s, were transplanted on June 7, 2018, with between-row spacing of 1.83 m and in-row spacing of 1.22 m. The same parental and F_1 plants were used as controls for all populations. Phenotypic data was collected three times a week for DFF. Leaf tissue was collected from all plants of the six populations and DNA was extracted and

KASP assays were performed to delineate the QTL region as previously described.

Validation using cultivars

A cultivar panel was compiled from watermelon of various pedigrees, consisting of both early and late flowering genotypes. The panel included the following: Sugar Baby (SB), NHM, Sugar Lump (SLMP), Strain II (SII), Mickylee (MICK), Golden Russian (GR), Charleston Gray (CG), Crimson Sweet (CS), AU-Producer (AUP), Calhoun Gray (CALG), Hopi Yellow (HOPI), Navajo Red (NAV), Sangria (SANG), KBS, Orangeglo (OG), Estrella (ES), and Allsweet (AS) (Supplementary Table S3).

In summer 2016 and 2017, 2 plants of each cultivar and the mapping population parents were transplanted per block in an RCBD with 5 blocks and 7 blocks, respectively. Data from the 2 years was averaged for analysis. All field experiments were done at the Durham Horticulture Farm in Watkinsville, GA. Similar to the RIL and recombinant populations, data was collected three times a week for DFF, after which leaf tissue was sampled for DNA extraction and genotyping as described earlier.

Results

SNP identification from candidate gene sequencing

The *Cla000855* (12,382,520–12,383,470 Mb (*TEM1*)) gene (951 bp) consists of 1 exon, while the *Cla009504* (14,874,778–14,879,348 Mb (*FT*)) gene (4,571 bp) consists of 4 exons (Guo et al. 2013). No SNPs were identified in the exons of either of these genes (data not shown). For *FT*, 8 SNPs were found within the intron (*Cla009504*_14876335, 14876592, 14877025, 14877848, 14877931, 14878293, 14878858, and 14878904). Sequencing 1,944 bp upstream of *FT* also revealed 5 SNPs between KBS and NHM. These SNPs include *Cla009504*_14872838, 14872843, 14872958, 14873014, and 14873563. *FT* was also sequenced 965 bp downstream but no SNPs were identified between KBS and NHM in this region. All the SNPs identified upstream and within the introns of *FT* had the same haplotype for all genotypes tested in this study. Only one SNP (*Cla009504*_14877931: KASP primer UGA3_14877931) is therefore shown to represent results from all *FT* SNPs (Supplementary Fig S1).

QTL-seq

A total of 206,385,328 (99.32%) and 215,194,945 (99.42%) mapped reads were obtained through QTL-seq for the E and L bulks, respectively. On chromosome 3, 38,240 SNPs were identified between the two bulks and a significant QTL was detected between ≈ 10.7 and 16.2 Mb (Fig. 1a), which overlapped with the region previously identified as the major QTL using classical QTL mapping (≈ 12–17 Mb) (McGregor et al. 2014). Several peaks were observed in this region, with 3 major peaks above the threshold signifying the possibility of more than one QTL/gene controlling the trait within the region. On further analysis of the significant regions, another potential candidate gene *Cla002795* (11,043,539–11,046,772 Mb) was identified at 11.04 Mb on chromosome 3 of watermelon. This gene is a phosphatidylinositol-4-phosphate 5-kinase (*PIP-kinase*), and is orthologous to a gene previously described to be involved in regulation of flowering time in *Arabidopsis* (Akhter et al. 2016) and rice (Ma et al. 2004). QTL-seq also identified numerous SNPs showing significant association with the flowering trait in this region. Twenty SNPs concentrated in the peak regions were selected for KASP assay development (Fig. 1b).

Marker performance in mapping population

KASP assays were designed for SNPs across the 10.7–16.2-Mb region (Supplementary Table S2) and tested on the KBS × NHM RIL population to determine which markers showed the highest association with the flowering phenotype. All markers within the QTL region tested on the RIL displayed significant association with the flowering trait ($P < 0.0001$), with R^2 values ranging from 0.37 to 0.63 (Supplementary Fig. S1). Based on the KASP assays, 22 recombinants out of the total 159 RILs were identified in the *Qdff3-1* region (Fig. 2a, Supplementary Fig. S1). The recombinants were used to identify regions in the QTL that displayed the highest association with the phenotype. Similar marker performance was observed from the 11.05–16.34-Mb region for most markers. However, three RILs, i.e., RIL-146 and RIL-112, which were early flowering, and RIL-4, which was late flowering, indicated that the region after NW0248748 (15.66 Mb) was

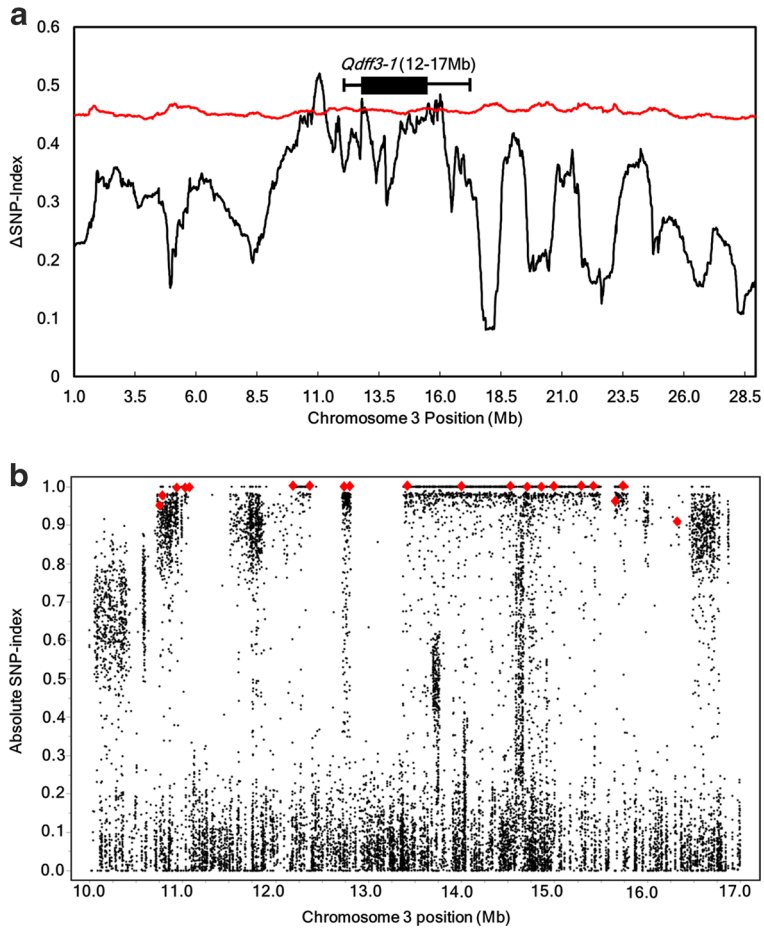

Fig. 1 **a** ΔSNP-index graph of the *Qdff3-1* locus (10.7–16.2 Mb) on chromosome 3 as validated by QTL-seq in the "Klondike Black Seeded" (late) × "New Hampshire Midget" (early) recombinant inbred line population plotted along with statistical confidence intervals under the null hypothesis of no QTL ($P = 0.01$) (red line). Box and whiskers above the peaks represent LOD-1 and LOD-1.5 support intervals, respectively, of the previously mapped locus using classical QTL mapping (McGregor et al. 2014). **b** SNPs converted to KASP assays (red diamonds). Megabase (Mb) positions are based on the 97103 watermelon genome (Guo et al. 2013)

not associated with the trait. Better marker association was observed from 11.05 to 15.66 Mb flanked by markers UGA3_11046548 and NW0248748, where most of the earlier and later flowering recombinants had the NHM and KBS allele respectively. This region hosts both the *TEM1* and *FT* genes. The QTL was therefore narrowed down to ≈ 4.62 Mb by genotyping the KBS × NHM RIL population with additional markers. One of the recombinants, RIL-190, was selected for additional fine-mapping purposes because it was heterozygous in the 11.05–15.66-Mb region, which showed the highest association with the trait (Supplementary Fig. S1).

Fine-mapping using recombinants

In addition to RIL-190, 376 F_2 plants were genotyped to identify more recombinants to further refine the QTL. Five recombinants, $KxNF_2$-204, $KxNF_2$-205, $KxNF_2$-217, $KxNF_2$-218, and $KxNF_2$-222, were selected from the F_2 plants following genotyping. These five F_2 recombinants were heterozygous in different regions: from NW0247945 to UGA3_16339961, UGA3_10738714 to UGA3_14537958, UGA3_10738714 to UGA3_11046548, UGA3_11046548 to UGA3_15774973, and UGA3_11046548 to UGA3_16339961, respectively (Fig. 2b). The RIL-190 (F_7) population (F_7RIL-190), which was heterozygous from UGA3_11046548 to UGA3_16339961, and the five $F_{2:3}$ populations were used for fine-mapping. The continuous phenotypic distributions of DFF in the six recombinant families confirmed the quantitative nature of the trait (Fig. 3). The average DFF was 16.17 for the earliest flowering population (KxN3-205) and 27.22 for the latest flowering population (KxN3-217). F_7RIL-190, KxN3-204, KxN3-218, and KxN3-222 populations had

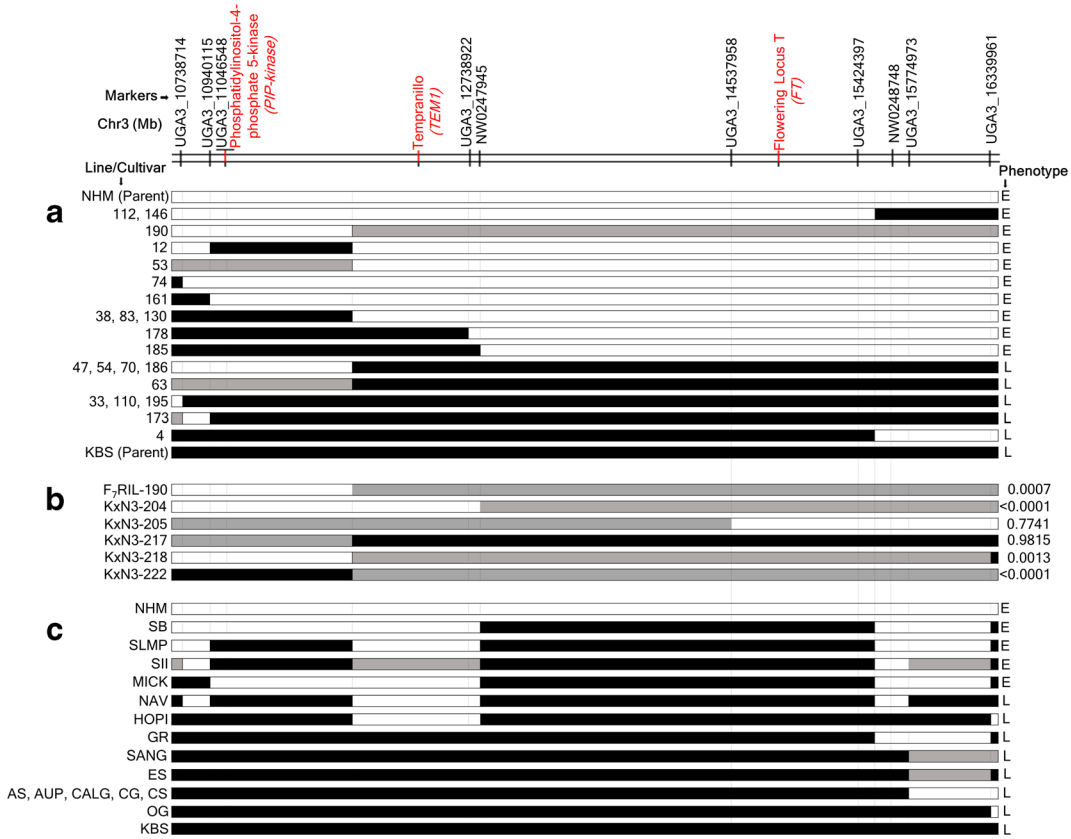

Fig. 2 Fine-mapping of *Qdff3-1* locus using (a) recombinants from the KBS × NHM RIL mapping population, (b) populations used for fine-mapping, and (c) cultivar panel. The physical map with markers are shown at the top, while the black, white, and gray bars denote the marker genotypes of KBS, NHM, and heterozygous alleles, respectively. On the right is the phenotype of the specific line (E = early and L = late) and the significance level associated with progeny testing of the recombinant populations. On the left is the specific line or cultivar identification

average DFFs of 25.84, 21.13, 25.55, and 18.91, respectively.

Progeny testing of the six families showed that the QTL was located between UGA3_14537958 and UGA3_16339931. A significant difference between the NHM-type and KBS-type progeny was observed in the segregating regions of all populations except KxN3-205 and KxN3-217 (Table 1; Fig. 2b). These two populations, which were the earliest and latest flowering, respectively, were both homozygous between UGA3_14537958 and UGA3_16339931 markers. Based on results from the KBS × NHM RIL mapping population, with NW0248748 already determined as the right flanking marker, the QTL was further narrowed down to ≈ 1.13 Mb with results from the six populations (Fig. 2b). The region is flanked by markers UGA3_14537958 and NW0248748 and includes *FT* (*Cla009504*).

Cultivar panel

The majority of the cultivar panel had an intermediate to later flowering time, with few earlier flowering cultivars. Average DFF of the cultivar panel for 2016 and 2017 ranged from 16.4 to 36.8, with SB and AS being the earliest and latest flowering cultivars, respectively. SB, NHM, SLMP, SII, and MICK were designated as "early" cultivars (DFF = 16.4-24.6) while GR, CG, CS, AUP, CALG, HOPI, NAV, SANG, KBS, OG, ES, and AS (DFF = 30.2–36.8) were in the "late" category (Supplementary Fig. S2).

Marker performance was tested on the panel with different genetic backgrounds (Fig. 2c). Three markers (UGA3_10738714, UGA3_10795402, and UGA3_11016809) in the 10.7- to 11.04-Mb region were highly associated (R^2 = 0.80) with the

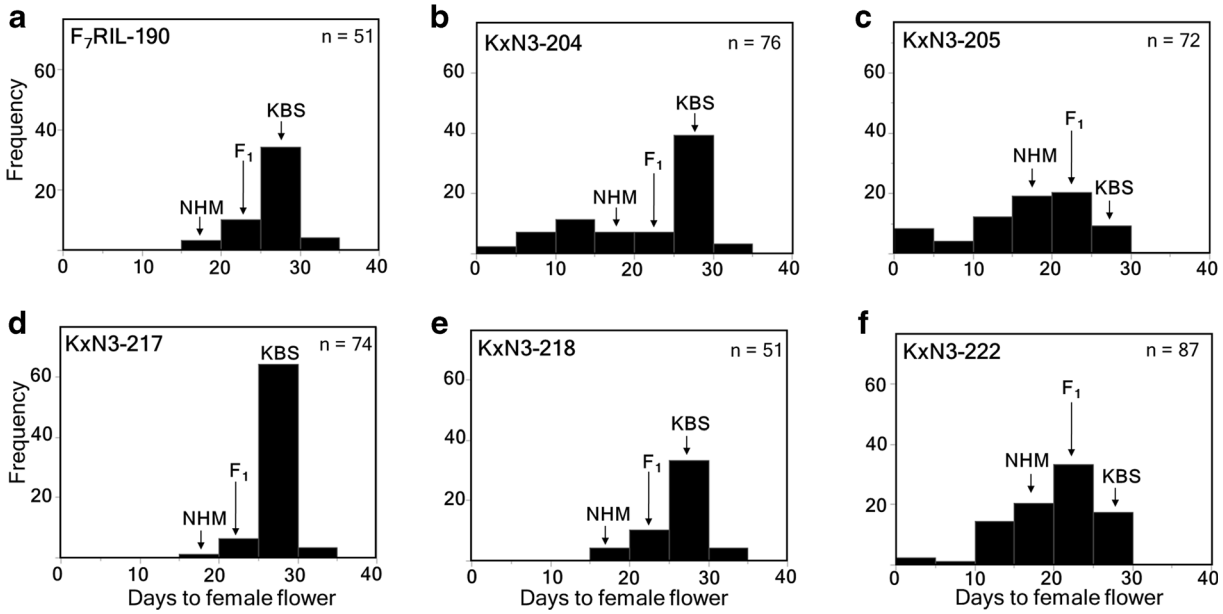

Fig. 3 Frequency distribution of flowering time in six (5 $F_{2:3}$ and 1 F_7 RIL) watermelon populations. **a** F_7RIL-190, **b** KxN3-204, **c** KxN3-205, **d** KxN3-217, **e** KxN3-218, **f** KxN3-222. The average DFF of the different populations was 25.84, 21.13, 16.17, 27.22, 25.55, and 18.91 days, respectively

phenotype (Supplementary Fig. S2). This is in contrast to the KBS × NHM RIL population, where fine-mapping excluded this region (Fig. 2), suggesting an additional region associated with control of flowering time in these diverse genetic backgrounds. The NW0248748 marker (15.66 Mb; R^2 = 0.60) gave the expected genotype for all cultivars except Golden Russian and Navajo. This marker was also the right flanking marker identified from the F_6 KBS × NHM RIL population recombinants, further validating the association of this region with DFF. Taking into account the cultivar panel, the 239-kb region between UGA3_15424397 and NW0248748 markers (Fig.

Table 1 Progeny test of recombinant plants

Parent	Progeny	Flowering time progeny test								
		KBS[a]				NHM[b]				
		n	Mean μ	Variance σ^2	s.d σ	*n*	Mean μ	Variance σ^2	s.d σ	*P* value
F_6RIL-190	F_7RIL-190	16	28.38	7.32	2.70	9	22.89	10.56	3.26	0.0007
KxN2-204	KxN3-204	23	28.17	13.51	3.68	14	11.57	61.80	7.86	< 0.0001
KxN2-205	KxN3-205	14	16.07	53.61	7.32	19	14.32	65.78	8.11	0.7741
KxN2-217	KxN3-217	17	27.88	3.74	1.93	13	27.69	6.90	2.63	0.9815
KxN2-218	KxN3-218	16	28.38	3.98	2.00	12	23.17	18.56	4.30	0.0013
KxN2-222	KxN3-222	21	24.12	23.69	4.87	24	17.33	38.39	6.20	< 0.0001

n = number of watermelon plants, *Mean* = average days to female flower, *s.d* standard deviation

[a] Average DFF for KBS-type progeny

[b] Average DFF for NHM-type progeny

2c) displayed significant association with the trait. There are 15 genes (Supplementary Table S4) within this region, 7 with unknown function.

Marker Performance

UGA3_15424397 and NW0248748 showed higher association in the mapping and recombinant populations compared to UGA3_10795402. However, in the cultivar panel, UGA3_15424397 did not exhibit high association with the trait as all cultivars, except NHM, had the KBS genotype. The NW0248748 marker displayed better performance than UGA3_15424397 with all the earlier flowering cultivars having the NHM genotype while the later flowering cultivars had the KBS genotype, except Golden Russian and Navajo. UGA3_10795402, unlike what was observed on the populations, exhibited significantly higher association with the trait on the cultivar panel (Fig. 4).

Discussion

Flowering time plays a key role in watermelon production as it is a major determinant of earliness and also a major factor in seedless watermelon production. Genetic control of flowering time is generally quantitative in nature and displays a continuous phenotypic variation. Several genes are involved in the regulation of this trait and flowering locus T (*FT*) has been described as one of the major genes that controls flowering in several species (Ahn et al. 2006; Kardailsky et al. 1999; Kobayashi et al. 1999; Lin et al. 2007; Lu et al. 2014; Schwartz et al. 2009). In this study, we used recombinants from a KBS × NHM RIL mapping population, KBS × NHM $F_{2:3}$ families, and a cultivar panel of different watermelon genotypes to refine the *Qdff3-1* locus controlling flowering time in watermelon.

Previously, a major QTL (≈ 12–17 Mb) governing flowering time was identified on chromosome 3 of watermelon using conventional QTL mapping (McGregor et al. 2014). Among the 172 predicted genes within this region, *FT* and *TEM1* were previously proposed as candidate genes controlling flowering time in watermelon (McGregor et al, 2014). These two genes were sequenced in this study but no SNPs were identified in the exons of either of the genes. *FT*, which was a more obvious candidate gene, was also sequenced 1,944 bp upstream and 965 bp downstream. Since the QTL region described in McGregor et al. (2014) was very large (≈ 5 Mb; 3.8 cM), we wanted to narrow down the region to identify a candidate gene and SNPs that could be useful for marker-assisted selection. A significant QTL was observed on chromosome 3 of watermelon from 10.7 to 16.2 Mb, validating the previous results. QTL-seq has been found to be highly effective in identifying markers linked to quantitative traits of agronomic interest in different crops including rice, cucumber, chickpea, and peanut (Clevenger et al. 2018; Das et al. 2015; Illa-Berenguer et al. 2015; Lu et al. 2014; Takagi et al. 2013). More recently, it has been used to identify a dwarfism gene in watermelon (Dong et al. 2018). Several SNPs were identified in this region and were converted to KASP assays for high-throughput genotyping. The region was fine-mapped using KBS × NHM populations developed from recombinants, which refined the region to a 1.13-Mb interval flanked by markers UGA3_14537958 and NW0248748. The *FT* gene (*Cla009504)* is located within this region.

Additional recombinants to further delineate the region were not available in the KBS × NHM genetic background. We turned our attention to historical recombination found in watermelon cultivars. All the cultivars in the cultivar panel, including early cultivars, had the late (KBS) genotypes in the region spanning the *FT* gene (UGA3_14537968 to UGA3_15424397). *FT* transcription in *Arabidopsis* is controlled by distant (4.0–5.7 kb) *cis*-regulatory elements upstream of *FT* (Adrian et al. 2010), but the KBS haplotype block in the cultivar panel stretches over ~ 2 Mb (UGA3_13418651 to UGA3_15424397). However, we only investigated SNPs and many other polymorphisms and modifications can affect gene function. Therefore, we cannot totally exclude the *FT* gene as playing a role in flowering time in these genetic backgrounds.

Marker NW0248748 had the expected allele for all but 2 of the cultivars. A possible explanation for our results is that the gene associated with DFF is found in the 239 kb between UGA3_15424397 and NW0248748. Recombination on either side of the gene, between the flanking markers would give the observed results. Despite several attempts, we were unable to design additional KASP assays within this region to determine more accurately the recombination point. The 239-kb region contains 15 genes, including *Cla009546*, which is a protein phosphatase 2C (*PP2C*). *PP2C* has been described as a positive regulator of

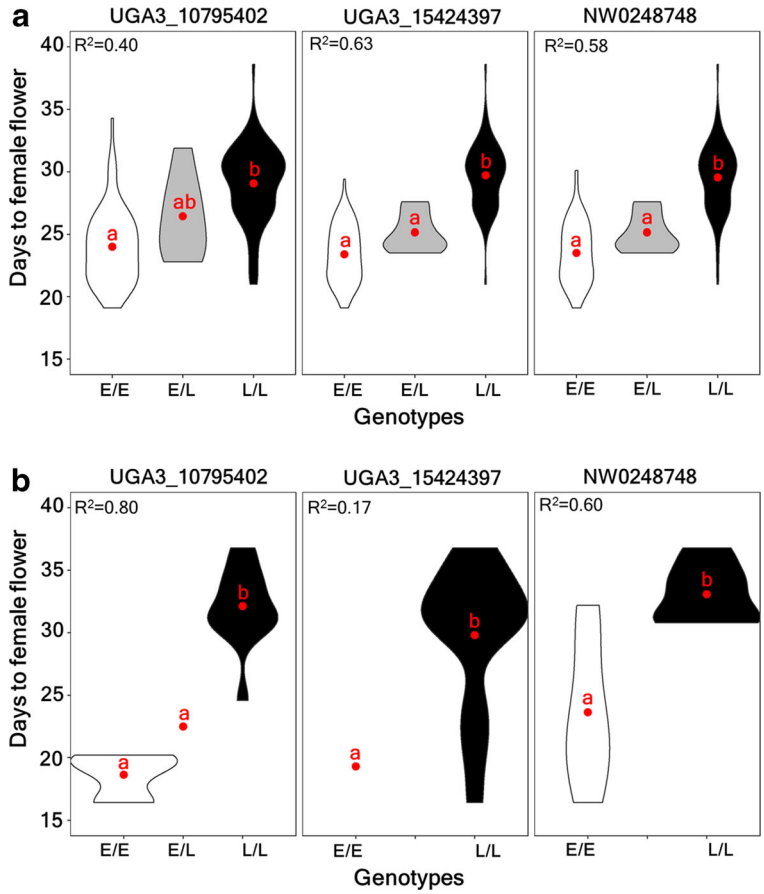

Fig. 4 Performance of UGA3_10795402, UGA3_15424397, and NW0248748 on **a** RIL mapping population and **b** cultivar panel. Dots indicate mean and levels not connected by the same letter are significantly different. E/E homozygous NHM (early) type allele, L/L homozygous KBS (late) type allele, and E/L heterozygous

flowering through regulation of the transcript levels of integrators and floral meristem identity genes, such as *Flowering locus C* (*FLC*), *CO*, *Suppressor of overexpression of CONSTANS 1* (*SOC1*), *LEAFY* (*LFY*), *FT* and *Flowering Locus D* (*FD*) in *Arabidopsis* (Zhai et al. 2016) (Supplementary Table S3). This study proposes it as a major candidate gene that contributes to the genetic variation in watermelon flowering time. Further research quantifying gene expression of these candidate genes in diverse watermelon cultivars will be needed to confirm the roles of the different candidate genes in the control of flowering time in watermelon.

High associations (R^2 = 0.80) were observed for markers between UGA3_10738714 and UGA3_11016809 and DFF in the cultivar panel. This was contrary to observations in the KBS × NHM RIL and recombinant populations, where markers in this region displayed low association with the trait (0.37–0.45) and progeny tests found no significant difference in DFF between KBS and NHM genotypes. These results suggest that there is differential regulation of flowering in the diverse genetic backgrounds represented in the cultivar panel.

The most appropriate marker to use for marker-assisted selection for flowering time in watermelon will depend on the genetic background used. UGA3_15424397 is effective for selection for DFF in the KBS × NHM background, but does not perform well in other backgrounds. NW0248748 may be useful to select for early flowering when SB (red flesh), SLMP (white flesh), SII (yellow flesh), or MICK (pink flesh) is used as a source of the early allele, as long as the non-donor is not NAV or GR. UGA3_10795402 might be useful when NAV or GR are used, as long as MICK is not the early flowering donor. The utility of these markers for marker-assisted selection in diverse backgrounds should be validated in segregating populations.

Conclusion

We refined the previously identified QTL region associated with DFF in KBS × NHM to a ≈ 1.13-Mb region of chromosome 3. A new candidate gene (*PP2C*) previously described to regulate flowering time genes in *Arabidopsis* was identified within this region. The results from diverse watermelon germplasm suggests that control of flowering time is dependent on genetic background. These findings provide more insight into the regulation of flowering time in watermelon and will have implications in watermelon breeding programs. The earlier harvesting days for the growers associated with the use of early flowering cultivars will reduce grower input costs and increase their profits.

Acknowledgments This work was in part supported by the United States Department of Agriculture (GEO00759) and the UGA Institute of Plant Breeding, Genetics, and Genomics.

References

Adrian J, Farrona S, Reimer JJ, Albani MC, Coupland G, Turck F (2010) cis-Regulatory elements and chromatin state coordinately control temporal and spatial expression of FLOWERING LOCUS T in Arabidopsis. Plant Cell 22(5): 1425–1440

Ahn JH, Miller D, Winter VJ, Banfield MJ, Lee JH, Yoo SY, Henz SR, Brady RL, Weigel D (2006) A divergent external loop confers antagonistic activity on floral regulators FT and TFL1. EMBO J 25(3):605–614

Akhter S, Uddin MN, Jeong IS, Kim DW, Liu X-M, Bahk JD (2016) Role of Arabidopsis AtPI4Kγ3, a type II phosphoinositide 4-kinase, in abiotic stress responses and floral transition. Plant Biotechnol J 14(1):215–230

Andrews, S (2010) FastQC: a quality control tool for high throughput sequence data. 8 Feb 2017. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bohuon EJR, Ramsay LD, Craft JA, Arthur AE, Marshall DF, Lydiate DJ, Kearsey MJ (1998) The association of flowering time quantitative trait loci with duplicated regions and candidate loci in Brassica oleracea. Genetics 150(1):393–401

Castillejo C, Pelaz S (2008) The balance between CONSTANS and TEMPRANILLO activities determines FT expression to trigger flowering. Curr Biol 18(17):1338–1343

Clevenger J, Chu Y, Chavarro C, Botton S, Culbreath A, Isleib TG, Holbrook CC, Ozias-Akins P (2018) Mapping late leaf spot resistance in peanut (Arachis hypogaea) using QTL-seq reveals markers for marker-assisted selection. Front Plant Sci 9:83

Das S, Upadhyaya HD, Bajaj D, Kujur A, Badoni S, Laxmi KV, Tripathi S, Gowda CL, Sharma S et al (2015) Deploying QTL-seq for rapid delineation of a potential candidate gene underlying major trait-associated QTL in chickpea. DNA Res 22(3):193–203

Dong W, Wu D, Li G, Wu D, Wang Z (2018) Next-generation sequencing from bulked segregant analysis identifies a dwarfism gene in watermelon. Sci Rep 8(1):2908

Ducrocq S, Giauffret C, Madur D, Combes V, Dumas F, Jouanne S, Coubriche D, Jamin P, Moreau L, Charcosset A (2009) Fine mapping and haplotype structure analysis of a major flowering time quantitative trait locus on maize chromosome 10. Genetics 183(4):1555–1563

Ferreira ME, Satagopan J, Yandell BS, Williams PH, Osborn TC (1995) Mapping loci controlling vernalization requirement and flowering time in Brassica napus. Theor Appl Genet 5(5):727–732

Guo S, Zhang J, Sun H, Salse J, Lucas WJ, Zhang H, Zheng Y, Mao L, Ren Y, Wang Z et al (2013) The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. Nat Genet 45(1):51–58

Illa-Berenguer E, Van Houten J, Huang Z, van der Knaap E (2015) Rapid and reliable identification of tomato fruit weight and locule number loci by QTL-seq. Theor Appl Genet 128(7): 1329–1342

Jimenez-Gomez JM, Alonso-Blanco C, Borja A, Anastasio G, Angosto T, Lozano R, Martinez-Zapater JM (2007) Quantitative genetic analysis of flowering time in tomato. Genome 50(3):303–315

Jung C, Müller AE (2009) Flowering time control and applications in plant breeding. Trends Plant Sci 14(10):563–573

Kardailsky I, Shukla VK, Ahn JH, Dagenais N, Christensen SK, Nguyen JT, Chory J, Harrison MJ, Weigel D (1999) Activation tagging of the floral inducer FT. Science 286(5446):1962–1965

King Z, Serrano J, Boerma HR, Li Z (2014) Non-toxic and efficient DNA extractions for soybean leaf and seed chips for high-throughput and large-scale genotyping. Biotechnol Lett 36(9):1875–1879

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(5446):1960–1962

Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M (2002) Hd3a, a rice ortholog of the Arabidopsis FT gene, promotes transition to flowering downstream of Hd1 under short-day conditions. Plant Cell Physiol 43(10):1096–1105

Levy YY, Dean C (1998) The transition to flowering. Plant Cell 10(12):1973–1989

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(16):2078–2079

Lifschitz E, Eviatar T, Rozman A, Shalit A, Goldshmidt A, Amsellem Z, Alvarez JP, Eshed Y (2006) The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. Proc Natl Acad Sci 103(16):6398–6403

Lin M-K, Belanger H, Lee Y-J, Varkonyi-Gasic E, Taoka K-I, Miura E, Xoconostle-Cázares B, Gendler K, Jorgensen RA, Phinney B et al (2007) Flowering locus T protein may act as

the long-distance florigenic signal in the cucurbits. Plant Cell 19(5):1488–1506

Lu H, Lin T, Klein J, Wang S, Qi J, Zhou Q, Sun J, Zhang Z, Weng Y, Huang S (2014) QTL-seq identifies an early flowering QTL located near flowering locus T in cucumber. Theor Appl Genet 127(7):1491–1499

Ma H, Xu S-P, Luo D, Xu Z-H, Xue H-W (2004) OsPIPK1, a rice phosphatidylinositol monophosphate kinase, regulates rice heading by modifying the expression of floral induction genes. Plant Mol Biol 54(2):295–310

Maheswaran M, Huang N, Sreerangasamy S, McCouch S (2000) Mapping quantitative trait loci associated with days to flowering and photoperiod sensitivity in rice (Oryza sativa L.). Mol Breed 6(2):145–155

Maynard DN, Elmstrom GW (1992) Triploid watermelon production practices and varieties. International Society for Horticultural Science (ISHS), Leuven, pp 169–178

McGregor CE, Waters V (2013) Pollen viability of F1 hybrids between watermelon cultivars and disease-resistant, infraspecific crop wild relatives. HortScience 48(12):1428–1432

McGregor CE, Waters V, Vashisth T, Abdel-Haleem H (2014) Flowering time in watermelon is associated with a major quantitative trait locus on chromosome 3. J Am Soc Hortic Sci 139(1):48–53

Milne I, Shaw P, Stephen G, Bayer M, Cardle L, Thomas WT, Flavell AJ, Marshall D (2010) Flapjack—graphical genotype visualization. Bioinformatics 26(24):3133–3134

Mohr HC (1986) Breeding vegetable crops. New york

Mouradov A, Cremer F, Coupland G (2002) Control of flowering time. Plant Cell 14(suppl 1):S111–S130

Ordonez Jr, SA, Silva, J, Oard, JH (2010) Association mapping of grain quality and flowering time in elite japonica rice germplasm. J Cereal Sci 51(3), 337-343

Osborn TC, Kole C, Parkin IAP, Sharpe AG, Kuiper M, Lydiate DJ, Trick M (1997) Comparison of flowering time genes in Brassica rapa, B. napus and Arabidopsis thaliana. Genetics 146(3):1123–1129

Poethig RS (2003) Phase change and the regulation of developmental timing in plants. Science 301(5631):334–336

Poland JA, Balint-Kurti PJ, Wisser RJ, Pratt RC, Nelson RJ (2009) Shades of gray: the world of quantitative disease resistance. Trends Plant Sci 14(1):21–29

Putterill J, Laurie R, Macknight R (2004) It's time to flower: the genetic control of flowering time. BioEssays 26(4):363–373

Rae AM, Howell EC, Kearsey MJ (1999) More QTL for flowering time revealed by substitution lines in Brassica oleracea. Heredity 83(5):586–596

Robert LS, Robson F, Sharpe A, Lydiate D, Coupland G (1998) Conserved structure and function of the Arabidopsis flowering time gene CONSTANS in Brassica napus. Plant Mol Biol 37(5):763–772

Sandlin K, Prothro J, Heesacker A, Khalilian N, Okashah R, Xiang W, Bachlava E, Caldwell DG, Taylor CA, Seymour DK (2012) Comparative mapping in watermelon [Citrullus lanatus (Thunb.) Matsum. et Nakai]. Theor Appl Genet 125(8):1603–1618

Schwartz C, Balasubramanian S, Warthmann N, Michael TP, Lempe J, Sureshkumar S, Kobayashi Y, Maloof JN, Borevitz JO, Chory J et al (2009) Cis regulatory changes at flowering locus T mediate natural variation in flowering responses of Arabidopsis thaliana. Genetics 183(2):723–732

Simpson GG, Dean C (2002) Arabidopsis, the Rosetta Stone of flowering time? Science 296(5566):285–289

Sumugat MR, Lee ON, Nemoto K, Sugiyama N (2010) Quantitative trait loci analysis of flowering-time-related traits in tomato. Sci Hortic 123(3):343–349

Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, Uemura A, Utsushi H, Tamiru M, Takuno S et al (2013) QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. Plant J 74(1):174–183

Thomson MJ, Tai TH, McClung AM, Lai X-H, Hinga ME, Lobos KB, Xu Y, Martinez CP, McCouch SR (2003) Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between Oryza rufipogon and the Oryza sativa cultivar Jefferson. Theor Appl Genet 107(3):479–493

Untergasser, A, Nijveen, H, Rao, X, Bisseling, T, Geurts, R, Leunissen, JA (2007) Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res 35(suppl2) W71-W74

U.S. Department of Agriculture (2017) U.S. Department of Agriculture- National Agricultural Statistics Service. Agricultural statistics. 10 Jan 2019. https://www.nass.usda.gov/Publications/Ag_Statistics/index.php

Wehner TC (2008) Watermelon. Vegetables I: Springer. p 381–418

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 103(51):19581–19586

Yano M, Kojima S, Takahashi Y, Lin H, Sasaki T (2001) Genetic control of flowering time in rice, a short-day plant. Plant Physiol 127(4):1425–1429

Zhai H, Ning W, Wu H, Zhang X, Lü S, Xia Z (2016) DNA-binding protein phosphatase AtDBP1 acts as a promoter of flowering in Arabidopsis. Planta 243(3):623–633

Zimmerman M (1988) Nectar production, flowering phenology, and strategies for pollination. Plant Reprod Ecol patterns Strateg 41:157–178