Theor Appl Genet (2014) 127:1491–1499
DOI 10.1007/s00122-014-2313-z

ORIGINAL PAPER

QTL-seq identifies an early flowering QTL located near *Flowering Locus T* in cucumber

Hongfeng Lu · Tao Lin · Joël Klein · Shenhao Wang · Jianjian Qi · Qian Zhou · Jinjing Sun · Zhonghua Zhang · Yiqun Weng · Sanwen Huang

Received: 26 December 2013 / Accepted: 10 April 2014 / Published online: 21 May 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract

Key message **Next-generation sequencing enabled a fast discovery of a major QTL controlling early flowering in cucumber, corresponding to the *FT* gene conditioning flowering time in *Arabidopsis*.**

Abstract Next-generation sequencing technologies are making it faster and more efficient to establish the association of agronomic traits with molecular markers or candidate genes, which is the requirement for marker-assisted selection in molecular breeding. Early flowering is an important agronomic trait in cucumber (*Cucumis sativus* L.), but the underlying genetic mechanism is unknown. In this study, we identified a candidate gene for early flowering QTL, *Ef1.1* through QTL-seq. Segregation analysis in F_2 and BC_1 populations derived from a cross between two inbred lines "Muromskij" (early flowering) and "9930" (late flowering) suggested quantitative nature of flowering time in cucumber. Genome-wide comparison of SNP profiles between the early and late-flowering bulks constructed from F_2 plants identified a major QTL, designated *Ef1.1* on cucumber chromosome 1 for early flowering in Muromskij, which was confirmed by microsatellite marker-based classical QTL mapping in the F_2 population. Joint QTL-seq and traditional QTL analysis delimited *Ef1.1* to an 890 kb genomic region. A cucumber gene, *Csa1G651710*, was identified in this region, which is a homolog of the *FLOWERING LOCUS T* (*FT*), the main flowering switch gene in *Arabidopsis*. Quantitative RT-PCR study of the expression level of *Csa1G651710* revealed significantly higher expression in early flowering genotypes. Data presented here provide support for *Csa1G651710* as a possible candidate gene for early flowering in the cucumber line Muromskij.

Communicated by M. J. Havey.

H. Lu and T. Lin contributed equally.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-014-2313-z) contains supplementary material, which is available to authorized users.

H. Lu · T. Lin · J. Klein · S. Wang · J. Qi · Q. Zhou · J. Sun · Z. Zhang · S. Huang (✉)
Key Laboratory of Biology and Genetic Improvement of Horticultural Crops of Ministry of Agriculture, Sino-Dutch Joint Lab of Horticultural Genomics, Institute of Vegetables and Flowers, CAAS, Beijing 100081, China
e-mail: huangsanwen@caas.cn

J. Klein
Laboratory of Molecular Biology, Department of Plant Science, Wageningen University, 6708 PB Wageningen, The Netherlands

Y. Weng
Vegetable Crops Research Unit, Department of Horticulture, United States Department of Agriculture (USDA), ARS, University of Wisconsin, Madison, WI, USA

Introduction

Proper timing of flowering has an important adaptive value for flowering plants. The transition from the vegetative phase to the reproductive phase is tightly controlled by multiple physiological signals and genetic pathways (Koornneef et al. 1998). In *Arabidopsis*, this transition is mainly under the control of *FLOWERING LOCUS T* (*FT*) which has been shown to encode a mobile floral signaling molecule (Kardailsky et al. 1999; Corbesier et al. 2007). *FT* homologs that play pivotal roles in floral induction have been identified in many crop plants such as tomato, *Solanum lycopersicum* L. (Lifschitz et al. 2006), rice, *Oryza sativa* L. (Tamaki et al. 2007), soybean, *Glycine max* (Kong et al. 2010), and onion, *Allium cepa* L. (Lee et al. 2013). In cucumber (*Cucumis*

sativus L.), one *FT* homolog is identified based on the sequence similarity with *FT* (Sato et al. 2009), but its relationship with flowering time in cucumber is not clear.

Flowering time is also an important trait and the target of selection in plant breeding. Genetic control of flowering time is in general quantitative in nature. Many studies have detected quantitative trait loci (QTL) related to flowering time or earliness in various crops. For example, in rice, 15 QTLs were associated with days to flowering in rice (Maheswaran et al. 2000). In tomato, three QTLs, *fast flowering* (*Ff*), *Ff-1* and *Ff-2*, were detected for earliness (Lindhout et al. 1994). Cucumber is the fourth most important vegetable worldwide (Tatlioglu 1993). Commercial cucumbers are in general photoperiod neutral. Early flowering is an important trait in cucumber that contributes to earliness and economic yield (Robbins and Staub 2009). Several studies identified QTLs for flowering time in cucumber. For example using an $F_{2:3}$ mapping population derived from a cross between cultivated and wild cucumber, Dijkhuizen and Staub (2002) detected two QTLs for days to anthesis with unknown chromosomal locations (one was presumably in chromosome 6 due to its linkage with the femaleness locus *F*). Using a RIL mapping population, Fazio et al. (2003) detected four QTLs controlling days to anthesis with low LOD support which were located in linkage groups 1,2,5, and 6, which seem to corresponded to chromosomes 6, 2, unknown, and 5, respectively (Weng et al. 2010). More recently, with a RIL population derived from two cultivated cucumber inbred lines 9930 and 9110Gt, Miao et al. (2012) identified a major effect QTL for days to anthesis for the first female flower, which was located in cucumber chromosome 1 flanked with microsatellite markers SSR22826 and SSR22638. However, the exact location of the QTL is not clear, not to mention the underlying genes and the precise genomic positions of these flowering time QTLs in cucumber are unknown.

Many agronomic traits in crop species such as flowering time, yield, and stress tolerance show a continuous phenotypic variation (Paterson et al. 1988). QTL mapping is the main approach for genetic dissection of quantitative traits, which provides the start point for map-based cloning of related genes and marker-assisted selection (MAS) in plant breeding. QTL mapping is usually conducted by genotyping a large number of individuals in segregating populations derived from bi-parental crosses, which is labor-intensive, time-consuming and sometimes costly (Salvi and Tuberosa 2005). The bulked-segregant analysis (BSA) (Michelmore et al. 1991) provides a simple and effective alternative to identify molecular markers linked to target genes or QTLs affecting a trait of interest by genotyping only a pair of bulked DNA samples from two sets of individuals with distinct or opposite extreme phenotypes. With the rapid development of next-generation sequencing (NGS) technologies, new strategies were proposed to take advantages of the power of BSA and NGS-aided high-throughput genotyping, which have been demonstrated in identifying major QTLs in yeast (*Saccharomyces cerevisiae*) (Ehrenreich et al. 2010; Swinnen et al. 2012), *Arabidopsis thaliana* (Schneeberger et al. 2009), rice (*Oryza sativa*. L) (Abe et al. 2012; Yang et al. 2013), and sunflower (*Helianthus annuus* L.) (Livaja et al. 2013). More recently, Takagi et al. (2013) described the QTL-seq approach for rapid mapping of quantitative trait loci in rice by whole genome re-sequencing of DNA bulks of phenotypic extremities. The objective of the present study was to identify major QTL for early flowering in cucumber using QTL-seq. We conducted whole genome re-sequencing of two DNA bulks, an early and a later flowering pool developed from plants in an F_2 population. Genome-wide SNP analysis allowed detection of a genomic region harboring the major early flowering QTL, which was confirmed with classical QTL analysis. Results from the study provided preliminary evidence that the cucumber *FT* gene is a possible candidate for this early flowering major QTL in cucumber.

Materials and methods

Plant materials and phenotyping for flowering time

Two cucumber inbred lines, "Muromskij" (CGN23617) and "9930" were used as parental lines to develop segregating populations for flowering time. Muromskij, from Russia, is one of the 115 core lines whose genome has been resequenced (Qi et al. 2013), and 9930 is a north China fresh market type cucumber (Chinese Long) whose draft genome assembly is available (Huang et al. 2009). Both lines are monoecious in sex expression, but Muromskij usually flowers 7–10 days earlier than 9930 after transplanting. Among 115 lines in the cucumber core collection we created before (Lv et al. 2012), Muromskij had the earliest flowering time. A cross was made between 9930 (female parent, P1) and Muromskij (pollen donor, P2) to create F_1, which was self-pollinated to generate the F_2 population, and backcrossed with 9930 to generate for BC_1P1 or with Muromskij for BC_1P2.

Flowering time of F_2 plants was recorded in three experiments conducted in 2012 spring, 2012 autumn and 2013 spring with 159, 232 and 258 F_2 individuals, respectively. Flowering time of the BC_1P1 (190 plants) populations was investigated only in spring of 2013. P1, P2 and F_1 plants (3–24 individuals) were included in all experiments. For each plant, the date of the first flower (either male or female) was recorded, and the days to anthesis after transplanting (DTA) was calculated as the flowering time of the plant. All the materials referred above were grown in the

greenhouses in Beijing, China and were under long daylight exposure during flowering period. The day and night average temperature of the greenhouse was controlled at 28 and 15 °C, respectively.

Generation and analysis of NGS data

Genome DNAs was isolated using the CATB method (Murray and Thompson 1980) from fresh leaves of the P1, P2, F_1 and F_2 plants from the 2012 autumn experiment which were used for both QTL-seq and SSR marker analysis.

For QTL-seq, two DNA pools, early pool (E-pool) and late pool (L-pool) were constructed, respectively, by mixing an equal amount of DNAs from 10 early flowering (DTA = 14–15) and 10 late flowering (DTA = 21–23) F_2 plants from the 2012 autumn experiment. Pair-end sequencing libraries (read length 100 bp) with insert sizes of around 500 bp were prepared for sequencing with an Illumina Genome Analyzer IIx machine.

The short reads from E-pool and L-pool were aligned to the 9930 reference genome (Huang et al. 2009) with the BWA software (Li and Durbin 2009). SNP-calling was performed by SAM tools software (Li and Durbin 2009). Low-quality SNPs with base quality value <20 and read depth <4× or those with >32× coverage from the E-pool sequences were excluded because these SNPs maybe false positives due to genomic repeat sequence, sequencing or alignment errors.

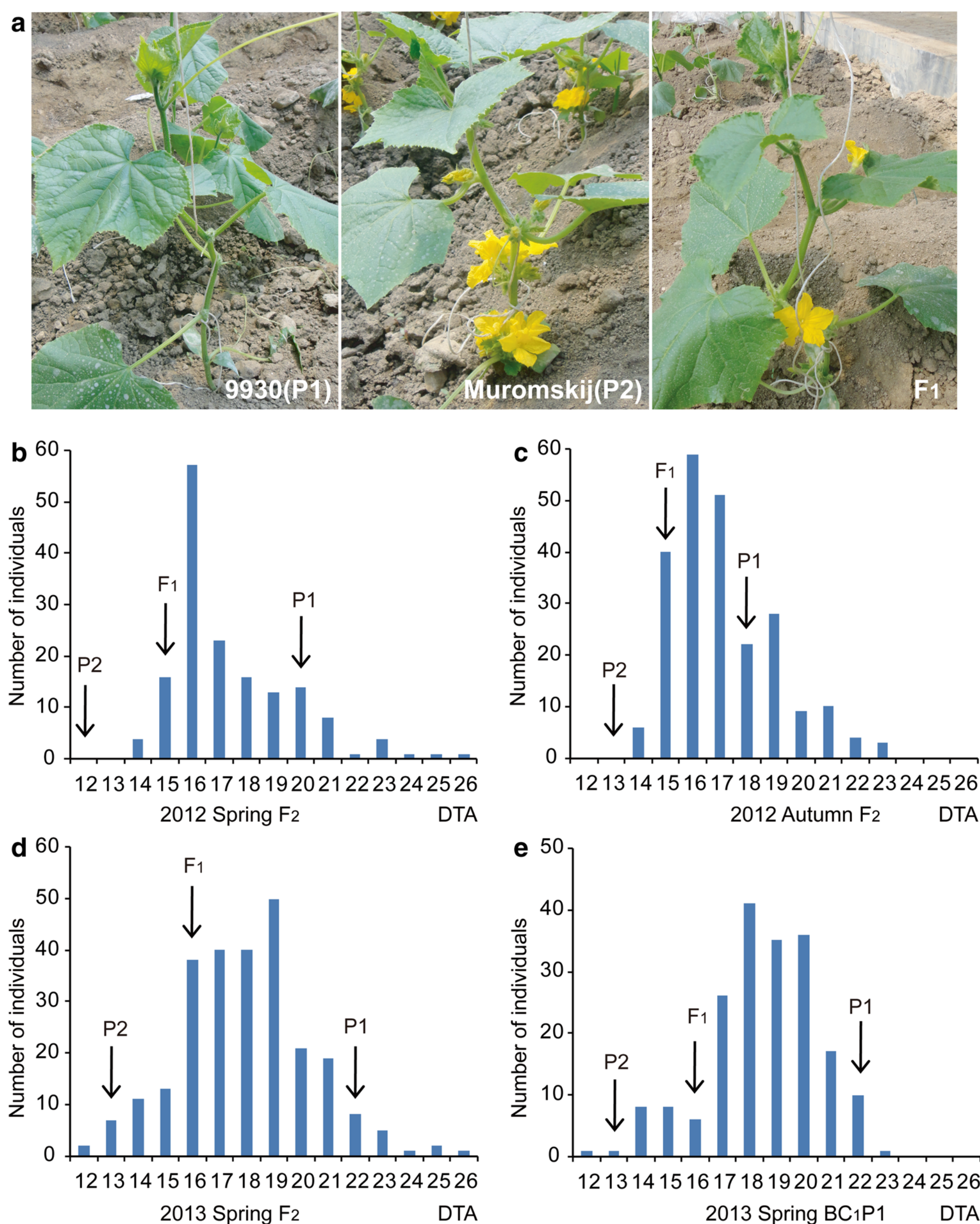

Fig. 1 Flowering time performance of two parents and their F_1, and frequency distribution of days to anthesis after transplanting among different populations in spring 2012, autumn 2012 and spring 2013 greenhouse experiments. **a** 9930 (P1, *left*), Muromskij (P2, *middle*) and their F_1 (*right*). Muromskij and F_1 had early flowering than 9930. Photos were taken 16 days after transplanting (Spring 2013). **b** The frequency of flowering time of P1, P2, F_1 and F_2 population in spring 2012. **c** The frequency of flowering time of P1, P2, F_1 and F_2 population in autumn 2012. **d–e** The frequency of flowering time of P1, P2, F_1, F_2 and BC_1P1 population in spring 2013

Two parameters, SNP-index and Δ (SNP-index) (Abe et al. 2012; Takagi et al. 2013) were calculated to identify candidate regions for early flowering QTL. An SNP-index is the proportion of reads harboring the SNP that are different from the reference sequence. Δ (SNP-index) was obtained by subtraction of SNP-index of E-pool from that of L-pool. Thus, SNP-index = 0 if the entire short reads contain genomic fragments from 9930; SNP-index = 1 if all the short reads were from Muromskij. An average of SNP-index of SNPs located in a given genomic interval was calculated using a sliding window analysis with 1 Mb window size and 10 kb increment. The SNP-index graphs for E-pool and L-pools, as well as corresponding Δ (SNP-index) graph were plotted.

The Δ (SNP-index) value should not be significantly different from 0 in a genomic region no major QTL of the target gene (Takagi et al. 2013). We calculated statistical confidence intervals of Δ (SNP-index) for all the SNP positions with given read depths under the null hypothesis of no QTLs, and plotted them along with Δ (SNP-index). For each read depth, 95 % confidence intervals of Δ (SNP-index) were obtained following Takagi et al. (2013).

QTL analyses with SSR and Indel markers

The major QTL for early flowering identified from QTL-seq was verified with SSR marker-based traditional QTL analysis. SSR markers in the predicted region of cucumber chromosome 1 (Ren et al. 2009) were employed for polymorphism screening between the two parental lines, and between the E-pool and L-pool. Additional Indel (insertion or deletion) markers were identified by aligning E-pool Illumina reads to the 9930 reference genome with BWA/SAMtools software (Li and Durbin 2009). Primers for the Indel markers were designed with Primer 5 (http://www.PromerBiosoft.com). Polymorphic markers were applied to the F_2 population plants. Linkage analysis was performed with JoinMap 4.0 (Van Ooijen 2011). QTL analysis was conducted with MapQTL4.0 using the multiple QTL model (MQM mapping) procedure (Van Ooijen et al. 2002).

Expression analysis of flowering time candidate gene by real-time PCR

We investigated the expression pattern of *Csa1G651710*, which is a homolog of the *Arabidopsis FLOWERING LOCUS T* (*FT*) gene in cucumber using quantitative RT-PCR (qPCR). In the spring 2013 experiment, leaf samples at 15th node were collected from five P1, P2 and F_1 individuals 30 days after transplanting, respectively. At the same day, we selected the five early flowering individuals from BC_1P1 (B-E), five late-flowering individuals from BC_1P1 (B-L), five early flowering individuals from F_2 (F-E) and five late-flowering individuals form F_2 (F-L). Leaf samples were also collected at 15th node, respectively. Each sample collected was one repeat. So there were five biological repeats for P1, P2, F1, B-L, B-E, F-E and F-L. Total RNAs for all the samples were extracted with EasyPure Plant RNA Kit (TranGen Biotech, Beijing, China). Reverse transcription was conducted

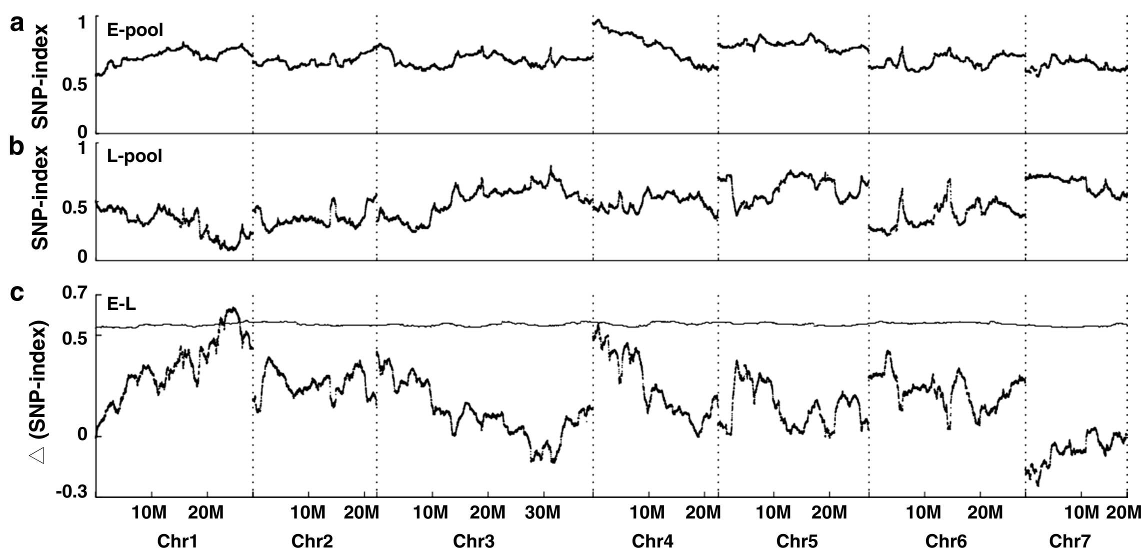

Fig. 2 SNP-index graphs of E-pool (**a**), L-pool (**b**) and Δ (SNP-index) graph (**c**) from QTL-seq analysis. *X*-axis represents the position of seven chromosomes and *Y*-axis represents the SNP-index. SNP-index was calculated based on 1 Mb interval with a 10 kb sliding window. The Δ (SNP-index) graph (**c**) was plotted with statistical confidence intervals under the null hypothesis of no QTL ($P < 0.05$). A candidate QTL (*Ef1.1*) location was identified in cucumber chromosome 1 (22.86–26.31 Mb interval) with the criteria that the SNP-index in E-pool (**a**) was near 1, SNP-index in L-pool (**b**) was near 0 and the Δ (SNP-index) (**c**) was above the confidence value ($P < 0.05$)

by Transcript one-step gDNA Removal and cDNA synthesis Supermix (TranGen Biotech). The qPCR primer pair sequences for *Csa1G651710* were 5′-TCGTGACCCTTT GGTTGTTGGGAGA-3′ (forward) and 5′-TCGGTCCCAC CAATCTCGACTCTTG-3′ (reverse). The cucumber β-actin gene (*Csa6G484600.1*) was used as an internal control (Li et al. 2012) which is a homolog to *ACTIN 7* in *Arabidopsis* (85 % CDS identity) (forward, 5′-ATTCTTGCATCTCTA AGTACCTTCC-3′ and reverse, 5′-CCAACTAAAGGGA AATAACTCACC-3′). *ACTIN7* was used as a positive control to ensure the quality of RNA and cDNA previously (Tanaka et al. 2005; Chambers and Shuai 2009). Each sample was repeated three times (technical replications). Average relative expression levels for P1, P2, F_1, B-L, B-E, F-E and F-L were calculated. *T* tests were performed to test the significance of differences in expression levels among different samples.

Results

Inheritance of early flowering time in Muromskij cucumber

Frequency distribution of flowering time among test materials in 2 years is presented in Fig. 1b–e. The average DTA (days to anthesis after transplanting) of Muromskij (P2) was about 5–9 days earlier than that of 9930 (P1), whereas F_1 plants flowered 2–3 days later than Muromskij and 3–6 days earlier than 9930 (Fig. 1) suggesting the major early flowering allele in Muromskij might be dominant. The DTAs in the four segregating populations (F_2-2012 Spring, F_2-2012 Autumn, F_2-2013 Spring, BC_1P1 in spring 2013) showed continuous variation, suggesting that the early flowering trait in Muromskijis quantitatively inherited (Fig. 1b–e).

QTL-seq identified Ef1.1, a major QTL locus controlling early flowering on chromosome 1

Illumina high-throughput sequencing resulted in 57,674,716 and 64,995,772 short reads (100 bp in length) from E-pool (8× depth coverage or 96.6 % coverage) and L-pool (9× depth coverage or 96.7 % coverage), respectively. These short reads were aligned to the 9930 reference genome and 234,393 SNPs were identified between E-pool and the reference genome. A SNP-index was calculated for each identified SNP. An average SNP-index was computed in a 1 Mb interval using a 10 kb sliding window. SNP-index graphs were generated for the E-pool (Fig. 2a) and L-pool (Fig. 2b) by plotting the average SNP-index against the position of each sliding window in the 9930 genome assembly. By combining the information of SNP-index in E-pool and L-pool, Δ (SNP-index) was calculated and plotted against the genome positions (Fig. 2c).

It was expected that the SNP-index graphs of the E-pool and L-pool would be identical for the genomic regions that are not relevant to the phenotypic difference (flowering time), whereas the genomic region(s) harboring the flowering time QTL(s) would exhibit unequal contributions from P1 and P2 parental genomes. In addition, the SNP-index of these regions for E- and L-pools would appear as mirror images with respect to the line of SNP-index = 0.5 (Takagi et al. 2013). The region on chromosome 1 from 22.86 to 26.31 Mb had an average SNP-index higher than 0.65 in E-pool with the highest being 0.84, while the SNP-index in the corresponding region of L-pool was lower than 0.31 (the lowest was 0.07). Examining SNP haplotypes among the early flowering individuals in the E-pool showed that most of those plants carried the Muromskij alleles in the 22.86–26.31 Mb region of chromosome 1, whereas most

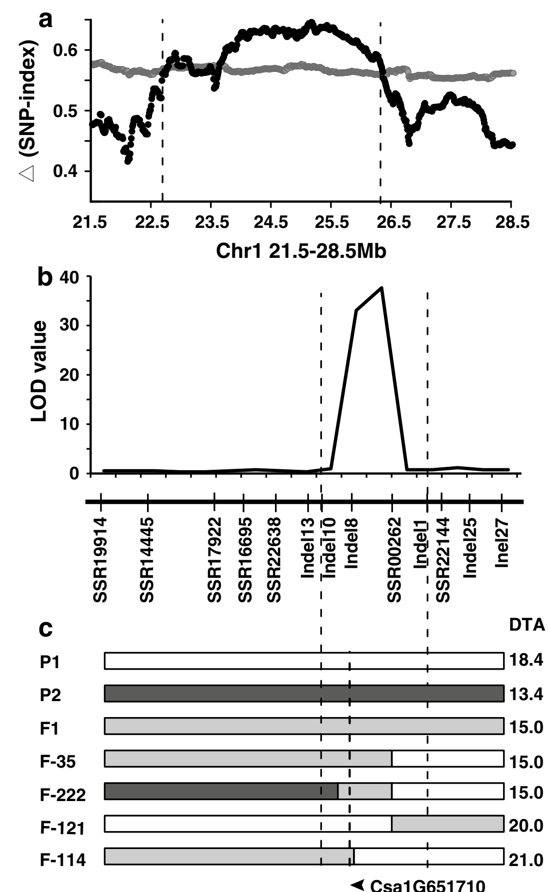

Fig. 3 Identification and validation of early flowering QTL *Ef1.1* in cucumber chromosome 1. **a** Δ (SNP-index) graph from QTL-seq analysis identified a QTL *Ef1.1* 1 at the interval of 22.86–26.31 Mb on chromosome 1. **b** Linkage analysis with molecular markers confirmed the location of *Ef1.1* QTL with closest flanking markers Inde10 to Indel1. **c** Examination of recombinants in F_2 refined the location of *Ef1.1* in an interval defined by Indel markers Indel8 and Indel1, where a candidate gene *Csa1G651710* was identified near SSR00262 (*arrowed*)

late-flowering individuals in the L-pool possessed 9930 alleles suggesting that there may be a major QTL controlling early flowering time in this region.

The Δ (SNP-index) value should be significantly different from 0 if a genomic region harbors a major QTL of the target gene. At 95 % significance level, only one genomic region on chromosome 1 from 22.86 to 26.31 Mb had the Δ (SNP-index) value that was significantly different from 0. These results indicated that there was a major QTL controlling flowering time at the 22.86–26.31 Mb region on chromosome 1 in cucumber (Fig. 3a), which was designated as *Ef1.1* (*early flowering 1.1*).

Analyses of the SSR and Indel markers narrowed down Ef1.1 to a 890 Kb interval

To confirm the early flowering QTL detected by QTL-seq, we conducted classical bi-parental QTL analysis with 232 F_2 plants from the 2012 autumn experiment. Among 142 SSR markers from chromosome 1 (Supplementary Table 1), seven were polymorphic between the E- and L-pools. From the distal part of chromosome 1 (21.0–29.1 Mb) which was not well covered by SSR markers, we developed six Indel markers. These 13 markers (details in Table 1) were applied to the segregating population for QTL analysis. MQM mapping analysis identified a major QTL for early flowering delimited by two Indel markers Indel10 and Indel1, which was physically located in the region of 25.42–27.31 Mb on chromosome 1 (Fig. 3b). The LOD scores in this region ranged from 0.04 to 37.3 (peaked at marker locus SSR00262), and this region could explain 52.3 % of the variance. This QTL mapping result was consistent with the QTL-seq analysis supporting a major QTL locus *Ef1.1* for early flowering in the genomic DNA interval of 25.42–26.31 Mb on chromosome 1. We analyzed SNP haplotypes of four F_2 recombinants in this region which allowed for further narrowing the *Ef1.1* locus down to an 1.84 Mb interval between the markers Indel8 (25.57 Mb) and Indel1 (27.31 Mb) (Fig. 3c). Therefore, the 25.42–26.31 Mb (890 kb) genomic regions in chromosome 1 may harbor the candidate gene for *Ef1.1*.

Identification of a candidate gene for the early flowering locus

In this 890 kb region, 84 genes were predicted (Li et al. 2011, Supplementary Table 2). Csa1G651710 caught our attention based on the gene annotation of cucumber which had the phosphatidylethanolamine-binding conserved site. Csa1G651710 was predicted to belong to the

Table 1 The information of 13 markers linked with early flowering on chromosome 1

Loci	Primer sequences (5′→3′)	Position (bp)[a]	LOD value
SSR19914	F: ATGGTCCACCAAACAAATGG R: GCTGTACTTGGAATCACTTCCC	21,334,709	0.18
SSR14445	F: TCCATGGAAATTGAAAACCC R: CGATCCTTATCGAACAGCCT	21,848,054	0.19
SSR17922	F: CATTCTAGGTCAATGAATCGCA R: GCAAAGTTGCCACATTGAAG	24,482,515	0.17
SSR16695	F: CACAATCCCACGAAGAACAA R: TGCAATTATGGCAAATCAAAA	24,649,281	0.36
SSR22638	F: TGTGTAAGATTTTTATTGGATGCC R: CTGAGCTTGATCAATTCCTTCA	24,682,912	0.21
Indel13	F: TGACACAAAAGTAACAAACATA R: GTCCATTCGTTAAGGAGTGA	25,262,753	0.04
Indel10	F: GCGATTGTAAAAGCATTG R: GCGATTGTAAAAGCATTG	25,426,904	0.60
Indel8	F: TAA AACACCACACCCGCAAT R: TT TAA AAGGGTATTGTTGG	25,570,083	32.73
SSR00262	F: CCGTTGGTCTTGGACTCTCA R: TGTAAAAGTGATCAGGAGGGTCT	25,851,547	37.25
Indel1	F: CTTCAAAGCGAAAAGGACGA R: ACATTACTCATTCTTGGCGA	27,308,465	0.42
SSR22144	F: AGGCTTACAGAACAGCATTA R: GCTGAGGAACAATGGTAAAT	27,364,694	0.55
Indel25	F: GCAAAATACAACAAAGTAACCC R: TTCTTTCTGTGTTTGGTCTGTT	28,328,087	0.80
Indel27	F: TATTTA CTCATCGTATCATTTT R: AAGTATTTGTATATGGCTTTT	28,790,348	0.47

[a] Location in 9930 draft genome assembly (Huang et al. 2009)

phosphatidylethanolamine-binding protein (PEBP) family. FLOWERING LOCUS T (FT) in *Arabidopsis* was also a PEBP family member and was a major component of florigen that regulated flowering time in *Arabidopsis thaliana* (Taoka et al. 2013). The results of blast alignment showed that the CDS identity of *Csa1G651710* with *Arabidopsis FT* gene was as high as 74 % (Supplementary Fig. 1) and *Csa1G651710* encoded a protein sharing 78 % sequence identity with *Arabidopsis* FT protein (Fig. 4a). Since *Csa1G651710* was the only *FT* homolog in the cucumber genome and in the *Arabidopsis* genome *FT* was the best match of *Csa1G651710*, we therefore regarded *Csa1G651710* as the *FT* homolog in cucumber and designated this gene *CsFT*.

The physical position of the *CsFT* gene in cucumber genome is 25,850,971–25,855,507 bp in the 9930 draft genome assembly, which is very close to SSR00262 marker at which the LOD curve peaked for flowering time QTL *Ef1.1* (Fig. 3b). Therefore, *CsFT* is a most possible candidate gene for *Ef1.1*.

We investigated expression patterns of *CsFT* with qPCR in two parental lines, their F_1, and the early and late-flowering individuals in F_2 and BC_1 populations to analysis if the expression level of *CsFT* may contribute to the variance of flowering time (Fig. 4b). The expression level of *CsFT* in the early flowering P2, F_1, F_2 (F-E) and BC_1 (B-E) plants was significantly higher than that in the late-flowering P1, F-L and B-L plants ($P < 0.05$) further suggesting that *CsFT* may be a candidate gene for the major QTL controlling flowering time in Muromskij cucumber.

Discussion

In this study, we employed QTL-seq (Takagi et al. 2013) to identify a major QTL for early flowering in cucumber with an F_2 mapping population. This method took advantage of the high-throughput whole genome re-sequencing and bulked-segregant analysis (BSA). In addition, use of SNP-index allowed accurate quantitative evaluation of the frequencies of parental alleles as well as the genomic contribution from the two parents to F_2 individuals. These features of QTL-seq make it a quicker and more efficient method to identify genomic regions harboring the major QTL of the target gene.

QTL-seq analysis identified a major QTL *Ef1.1*, on cucumber chromosome 1, which was verified with classical QTL analysis (Fig. 3). The position of *Ef1.1* was defined by two flanking SSR markers SSR22638 and SSR22144 (24.68–27.36 Mb). *Ef1.1* is near to QTL *Da1.1/1.2*, controlling for the first female flowering time detected by Miao et al. (2012). *Da1.1/1.2* located in between 23.72–24.68 Mb on cucumber chromosome 1, defined by two flanking markers SSR22826 and SSR22638. Therefore, *Da1.1/1.2* and *Ef1.1* are close but different QTLs.

By a closer look of the SNP-index and Δ (SNP-index) plots (Fig. 2), in addition to *Ef1.1*, two regions also showed slight deviations from 0 in Δ (SNP-index) values: one was an interval of 27.61–33.62 Mb on chromosome 3, and the other on almost the entire chromosome 7. Both regions had negative Δ (SNP-index) values (Fig. 2) as opposed to what was observed at the *Ef1.1* region on chromosome 1. We did not find the polymorphic SSR markers on chromosome 7 through screening the 108 SSR markers on chromosome 7 and the linkage analysis on chromosome 3 have not been performed by now. This may suggest that the early flowering parental line Muromskij might also bear late-flowering

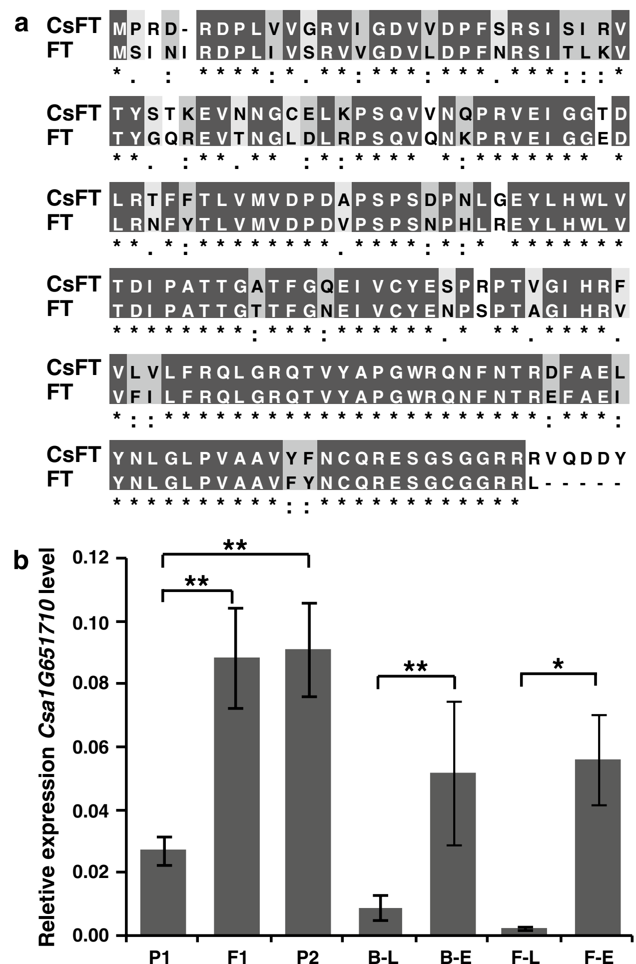

Fig. 4 Structure identity and expression of cucumber *FT* gene homolog *Csa1G651710*. **a** Alignment of cucumber Csa1G651710 and *Arabidopsis* FLOWERING LOCUS T (FT) protein sequence. Amino acid residues with >50 % identity or similarity between the two homologs are *shaded black* or *gray*, respectively. **b** Relative expression of *Csa1G651710* in the early flowering Muromskij (P2), F_1, BC_1P2 (B-E) and F_2 (F-L) bulks was significant higher than later flowering 9930 (P1), BC_1P2 (B-L) and F_2 (F-L) bulks. $^{**}P < 0.01$ and $^{*}P < 0.05$, respectively, in Student's *t* test. *Bar* SEM, $n = 5$

alleles, which is consistent with the transgressive segregation pattern observed in the segregating populations (Table 1). Further analysis is needed to identify those minor QTLs.

QTL-seq and classical QTL analysis delimited the early flowering QTL, *Ef1.1* to an 890 kb physical interval on chromosome 1 that contributed to 52.3 % of the phenotypic variation (Fig. 3). There are 84 predicted genes in this region. Among the 84 genes, the genetic position (Fig. 3) and expression pattern (Fig. 4b) of the *CsFT* gene suggested that it could be a candidate gene for *Ef1.1*. The physical location of *CsFT* was near to SSR00262 which had the highest LOD value (Fig. 4a). CsFT shared 78 % amino acid sequence identity with Arabidopsis FT (Fig. 4a) and the expression level of *CsFT* in early flowering plants was higher than the later flowering plants (Fig. 4b). Therefore, it is reasonable to postulate that *CsFT* is the candidate gene for the early flowering in cucumber. However, further evidence is needed to functionally validate this. For the remaining 83 genes in this region, there are 36 genes predicted to be unknown genes. For the 47 genes left, we did not identify any gene that is homologous *Arabidopsis* genes involved in the flowering time pathway, for example, *CO*, *SOC1*, *FD*, *TFL1* et al. However, we need to further experiments to prove the hypothesis.

Acknowledgments This work was supported by funding from the National Natural Science Foundation of China (NSFC: 31225025), the National Program on Key Basic Research Projects in China (The 973 Program: 2012CB113900), the National High Tech Research Development Program in China (The 863 Program: 2010AA10A108, 2012AA100101), and the Chinese Ministry of Finance (1251610601001).

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standard The experiments in this study comply with the current laws of China.

References

Abe A, Kosugi S, Yoshida K, Natsume S, Takagi H, Kanzaki H, Matsumura H, Yoshida K, Mitsuoka C, Tamiru M (2012) Genome sequencing reveals agronomically important loci in rice using MutMap. Nature Biotech 30:174–178

Chambers C, Shuai B (2009) Profiling microRNA expression in *Arabidopsis* pollen using microRNA array and real-time PCR. BMC Plant Biol 9(1):87

Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C (2007) FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. Science 316:1030–1033

Dijkhuizen A, Staub JE (2002) QTL conditioning yield and fruit quality traits in cucumber (*Cucumis sativus* L.) effects of environment and genetic background. J New Seeds 4:1–30

Ehrenreich IM, Torabi N, Jia Y, Kent J, Martis S et al (2010) Dissection of genetically complex traits with extremely large pools of yeast segregants. Nature 464:1039–1042

Fazio F, Staub JE, Stevens MR (2003) Genetic mapping and QTL analysis of horticultural traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. Theor Appl Genet 107:864–874

Huang S, Li R, Zhang Z, Li L, Gu X, Fan W, Lucas WJ, Wang X, Xie B, Ni P (2009) The genome of the cucumber, *Cucumis sativus* L. Nature Genet 41:1275–1281

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

Kong F, Liu B, Xia Z, Sato S, Kim B, Watanabe S, Yamada T, Tabata S, Kanazawa A, Harada K, Abe J (2010) Two coordinately regulated homologs of *FLOWERING LOCUS T* are involved in the control of photoperiodic flowering in soybean. Plant Physiol 154:1220–1231

Koornneef M, Alonso-Blanco C, Peeters AJ, Soppe W (1998) Genetic control of flowering time in *Arabidopsis*. Annu Rev Plant Biol 49:345–370

Lee R, Baldwin S, Kenel F, McCallum J, Macknight R (2013) *FLOWERING LOCUS T* genes control onion bulb formation and flowering. Nat Commun 4:2884

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760

Li Z, Zhang ZH, Yan PC, Huang SW, Fei ZJ, Lin K (2011) RNA-seq improves annotation of protein-coding genes in the cucumber genome. BMC Genom 12:540

Li Q, Zhang C, Li J, Wang L, Ren Z (2012) Genome-wide identification and characterization of *R2R3MYB* family in *Cucumis sativus*. PLoS One 7(10):e47576

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 USA 103:6398–6403

Lindhout P, Van Heusden S, Pet G, Van Ooijen JW, Sandbrink H, Verkerk R, Vrielink R, Zabel P (1994) Perspectives of molecular marker assisted breeding for earliness in tomato. Euphytica 79:279–286

Livaja M, Wang Y, Wieckhorst S, Haseneyer G, Seidel M, Hahn V, Knapp SJ, Taudien S, Schön C-C, Bauer E (2013) BSTA: a targeted approach combines bulked segregant analysis with next-generation sequencing and de novo transcriptome assembly for SNP discovery in sunflower. BMC Genom 14:628

Lv J, Qi J, Shi Q, Shen D, Zhang S, Shao G, Li H, Sun Z, Weng Y, Shang Y (2012) Genetic diversity and population structure of cucumber (*Cucumis sativus* L.). PLoS One 7:e46919

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 6:145–155

Miao H, Gu XF, Zhang SP, Zhang ZH, Huang SW, Wang Y, Fang ZY (2012) Mapping QTLs for seedling-associated traits in cucumber. Acta Hort Sinica 39:879–887

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

Murray M, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8:4321–4326

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721–726

Qi J, Liu X, Shen D, Miao H, Xie B, Li X, Zeng P, Wang S, Shang Y, Gu X (2013) A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. Nat Genet 45:1510–1515

Ren Y, Zhang Z, Liu J, Staub JE, Han Y, Cheng Z, Li X, Lu J, Miao H, Kang H (2009) An integrated genetic and cytogenetic map of the cucumber genome. PLoS One 4:e5795

Robbins MD, Staub JE (2009) Comparative analysis of marker-assisted and phenotypic selection for yield components in cucumber. Theor Appl Genet 119:621–634

Salvi S, Tuberosa R (2005) To clone or not to clone plant QTLs: present and future challenges. Trends Plant Sci 10:297–304

Sato H, Heang D, Sassa H, Koba T (2009) Identification and characterization of *FT/TFL1* gene family in cucumber. Breed Sci 59:3–11

Schneeberger K, Ossowski S, Lanz C, Juul T, Petersen AH, Nielsen KL, Jørgensen J-E, Weigel D, Andersen SU (2009) SHOREmap: simultaneous mapping and mutation identification by deep sequencing. Nat Methods 6:550–551

Swinnen S, Schaerlaekens K, Pais T, Claesen J, Hubmann G et al (2012) Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis. Genome Res 22:975–984

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

Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K (2007) Hd3a protein is a mobile flowering signal in rice. Science 316:1033–1036

Tanaka Y, Sano T, Tamaoki M, Nakajima N, Kondo N, Hasezawa S (2005) Ethylene inhibits abscisic acid-induced stomatal closure in Arabidopsis. Plant Physiol 138(4):2337–2343

Taoka K, Ohki I, Tsuji H, Kojima C, Shimamoto K (2013) Structure and function of florigen and the receptor complex. Trends Plant Sci 18:287–294

Tatlioglu T (1993) Cucumber *Cucumis sativus* L. Genetic improvement of vegetable crops. Pergamon Press, Tarrytown, pp 197–234

Van Ooijen JW (2011) Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. Genet Res 93(5):343–349

Van Ooijen JW, Boer MP, Jansen RC, Maliepaard C (2002) MapQTL4.0, software for the calculation of QTL positions on genetic maps. Plant Research International, Wagenigen

Weng Y, Johnson S, Staub JE, Huang SW (2010) An extended microsatellite genetic map of cucumber, *Cucumis sativus* L. HortScience 45:880–886

Yang Z, Huang D, Tang W, Zheng Y, Liang K, Cutler AJ, Wu WR (2013) Mapping of quantitative trait loci underlying cold tolerance in rice seedlings via high-throughput sequencing of pooled extremes. PLoS One 8(7):e68433